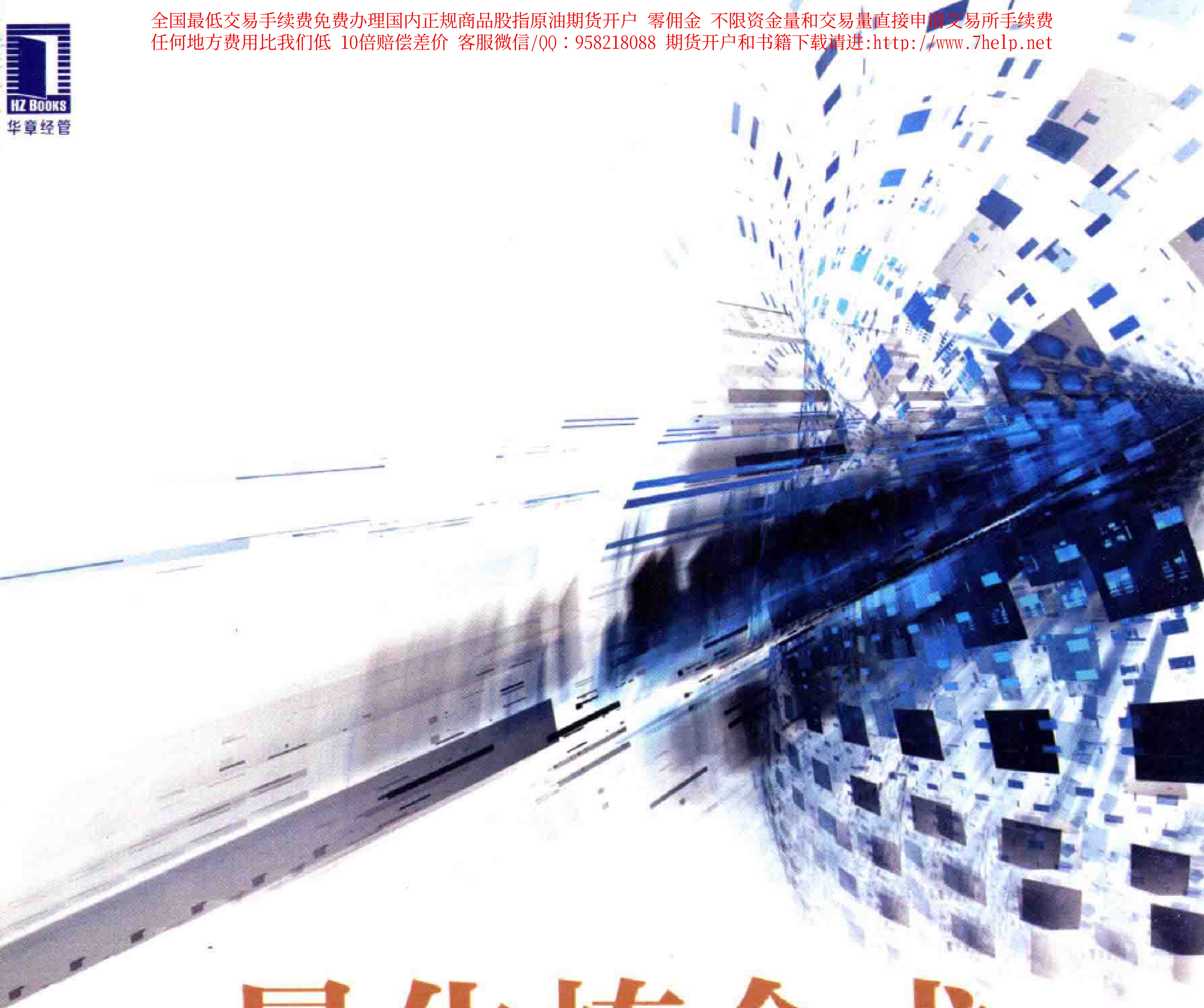
全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

HZ BOOKS
华章经管



量化炼金术

中低频量化交易策略研发

中低频量化交易领域的必备科学指南

通过中国金融市场上的实际案例，详细解构量化交易策略的研发流程

量化专业人士用心之作



机械工业出版社
China Machine Press

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

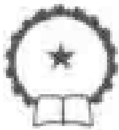


量化炼金术

中低频量化交易策略研发

杨博理 贾芳 著





机械工业出版社
China Machine Press

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书
www.55188.com
理想藏书55188

图书在版编目（CIP）数据

量化炼金术：中低频量化交易策略研发 / 杨博理，贾芳著．—北京：机械工业出版社，2017.8

ISBN 978-7-111-57532-0

I. 量… II. ① 杨… ② 贾… III. 股票交易 - 研究 IV. F830.91

中国版本图书馆 CIP 数据核字（2017）第 180901 号

本书首先对量化交易策略的历史、内涵和特点进行了简述，然后给出了量化交易策略的基本研发流程。书中介绍了策略研发中的一些注意事项与应对方法，随后通过两个例子说明了择时策略和选股策略的基础研发框架；在介绍推进分析这一回测技术的基础上，又进一步完善了择时策略和选股策略的研发流程。然后，作者增加了策略风险这一维度，并给出了论述。在此基础上，作者介绍了仓位优化的择时策略和投资组合优化的选股策略，从而使得策略有能力对整体风险进行控制；并在两种策略中增加了交易成本的计算，让量化交易策略的回溯测试更贴近现实交易的情况。针对这些完善的策略，本书给出了相应的策略评价。最后对整本书的内容进行了总结，并讨论了策略研发中的主观控制等相关问题。

量化炼金术：中低频量化交易策略研发

出版发行：机械工业出版社（北京市西城区百万庄大街 22 号 邮政编码：100037）
责任编辑：宋 燕　　责任校对：殷 虹
印　刷：北京瑞德印刷有限公司　　版　次：2017 年 8 月第 1 版第 1 次印刷
开　本：170mm×242mm 1/16　　印　张：15.25
书　号：ISBN 978-7-111-57532-0　　定　价：59.00 元

凡购本书，如有缺页、倒页、脱页，由本社发行部调换
客服热线：（010）68995261 88361066　　投稿热线：（010）88379007
购书热线：（010）68326294 88379649 68995259　　读者信箱：hzjg@hzbook.com

版权所有・侵权必究
封底无防伪标均为盗版
本书法律顾问：北京大成律师事务所 韩光 / 邹晓东

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

序言

接触过量化交易或者量化投资的人们，似乎常常会有这样一个执念，就是希望能找到一个万能的公式帮助他们击败市场、攫取利润。换句话说，是希望通过量化的手段获得交易的“圣杯”。而且，相较于技术分析而言，复杂的数学算法似乎是达成这样一个夙愿更有力的工具。

我自己很难判断这种万能的公式是否存在，并且以我有限的研究经验来看，找到万能的公式是一件十分困难的事情。当然，我依然对这件事情抱有希冀，谁没有个发财梦呢。不过经过这么多年的专业训练，我倒是有了另外一些有价值的收获。也正是这些收获的普适性和可分享性，驱使我写作了本书。

现在换一个角度来看待上面的问题。扪心自问，一个万能的盈利公式会被写入书籍资料中供人阅览吗？很明显，至少在量化交易这个领域内是没有可能的，策略可复制性太强，没有人会让庞大的读者群体瓜分自己的收益。即使存在讲解具体策略的书籍，也无外乎是一些已经不再具有显著盈利能力的交易策略，如海龟交易法则等。

这也是量化交易策略的另外一个问题，研发人员并不能保证一个策略在市场上一直有效。即便现在有效，交易策略能否适用于未来也是一个问题。如果交易策略的持续有效性很难保证，那“万能”“圣杯”便也无从谈起了。

那么在量化领域中，是否存在持续有效的事物呢？对“圣杯”的追求促使

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

我思考这个问题。当我尝试着跳出策略、公式的约束时，我似乎找到了一个可能的答案，那就是科学的研发行为和思维模式。

关于这一点，我自己是这样思考的：正向来看，量化交易策略是将交易思想通过数量化的规则进行表达；反向来看，量化交易策略的研发其实也是一种数据科学实践，只不过数据科学具体落脚在金融交易领域。既然是数据科学的具体实践，那么采用科学的方法、科学的思维就会是行之有效的手段，这是一件被过去许多年科学发展的历史所佐证的事情。

本书的写作过程，始终遵循了这样的理念，即以科学的方法构建量化交易策略、以科学的态度对待研发流程中的问题。例如，自回归、多因子模型、最优投资组合理论等都是经济金融领域的科学模型，回溯测试、推进分析、最优化等也可以看作数据科学方法在量化领域的运用与落地。而书中多次提及的数据结果与主观逻辑的结合，抽象出来其实就是归纳与演绎这两种科学方法的辩证统一。

需要声明的是，尽管我如此看重科学性在量化交易策略研发中的作用，但是有一点仍然应该认清，那就是学界和业界对于金融市场的科学研究目前来看还是不完善的，我自己对这一领域的掌握也不敢说成熟，因此也就不敢妄自尊大地将自己的方法论看作真正的“科学”。在本书中，作者仅仅是将自己关于研发量化交易策略的系统流程表述出来，希望对读者有所裨益。

这也是我将本书的书名取为《量化炼金术》的原因，除了向大师致敬以外，其实更重要的是想表达我对这个研究领域的敬畏。相较于真正的“科学”，也许“炼金术”才是我目前认知水平的实际体现。当然，暂且不论作者的水平如何，如果你能从本书中读出一些策略研发的科学精神，我也就足够欣慰了。

书中内容的安排大致如下：

第 1 章的引言部分介绍了量化交易策略的特性内涵与历史发展。比较遗憾的是，其发展历程主要都是对国外相关情况的介绍。作者也希望在未来，随着中国市场的成熟与发展，能有我们自己的研究、事件或公司可以丰富这部分内容。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

第 2 章是在具体介绍量化交易策略的研发之前，对整个研发流程和框架的总领性论述。第 3 章则是对研发量化交易策略的过程中可能会遇到的问题进行阐述，并且针对这些问题，说明哪些地方需要注意以及可以使用哪些方法进行处理。

第 4 ～ 14 章由浅入深地介绍了量化交易策略研发的具体流程和相关案例。量化交易策略大体上分为择时策略和选股策略两种，覆盖了时间序列和横截面两个维度。由浅入深则是指我们在介绍的过程中，逐步在简单策略中加入推进分析、对风险的考量、仓位决策等组成部分，从而使策略细节逐渐丰富，研究不断深化。

这一部分是本书的主体内容，我们将推进分析、收益、买卖、风险、仓位等量化交易策略会涉及的内容与组件，分拆开来进行介绍，再通过合适的实际案例逐步组合为一个整体。通过这样一种递进的方式对完整的策略研发过程进行讲解，能够帮助读者形成自己的研究思路，并使其体系化、流程化、逻辑化。我们希望读者在熟悉这一部分的内容之后，可以在所举案例的基础上举一反三，自行开发和检验适合自己的量化交易策略。

第 15 章是在完成了量化交易策略的研发工作之后，对具体策略给出评价报告。第 16 章包括两部分内容：一部分是全书内容的总结；另一部分则在于探讨所介绍的研发流程的局限性，同时给出了作者认为可行的应对方式，即主观逻辑的支持。

就全书而言，我们的着眼点主要是为量化交易策略的研发提供一个可以参考的流程框架。而具体的策略与建立在真实数据上的案例，虽然占据了书中大量的篇幅，但是其目的在于帮助读者更好地理解这个流程框架，因此策略本身也都是由最为简单的模型所构成的，本书并不介绍前沿的策略模型。

当然，我觉得读者是能够理解这样的处理方式的。正如前面所言，量化交易策略本身就是一个复制较为容易、不宜公开的策略种类。我们希望本书的读者能将更多的注意力放在对研发流程和研发思路的学习上，如果仅仅追求可盈利的具体策略，实践或许是一种比读书更为有效的途径。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

与此同时，为了使内容介绍更为清晰，书中在一些研究设置的细节处也基本上使用了最为简化的处理。例如，回溯测试中交易判断的频率和策略重新优化的频率，在实际研究中往往需要结合交易成本等因素分别加以设定；多因子模型之类的策略模型，也往往需要搭配更复杂的技术处理，从而贴近真实的交易环境。这些都是需要读者在具体实践时根据实际情况加以理解和把握的。

为了本书的最终完成，两位作者都投入了大量的心血和时间，但是限于水平和精力有限，书中难免存在着各种纰漏和问题。这里特别感谢北京大学魏鑫同学的仔细阅读，并指出上个版本中的一处收益率计算方式的偏误。也希望各位读者能在遇到疑难处时与作者沟通。出于一个研究人员的操守，本人将会认真对待书中的每一个问题，让本书能为读者提供更为精准、正确的指引。

杨博理

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

目录

序言

第 1 章 引言 ┊ 1

1.1 量化交易策略简述 ┊ 2
1.2 量化交易策略的优缺点 ┊ 8

第 2 章 量化交易策略的研发流程 ┊ 14

2.1 量化交易策略的基本研发流程 ┊ 15
2.2 量化交易策略研发流程的进一步论述 ┊ 18

第 3 章 注意事项与应对 ┊ 23

3.1 未来信息的规避 ┊ 24
3.2 过度拟合与欠拟合 ┊ 27
3.3 回溯测试与真实环境的差异 ┊ 31

第 4 章 简单的择时策略 ┊ 36

4.1 择时策略的基本框架 ┊ 37
4.2 均线趋势策略的简单优化 ┊ 40

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

4.3 均线反转策略的简单优化 ┊ 44
4.4 自回归策略的简单优化 ┊ 47

第 5 章 简单的选股策略 ┊ 51

5.1 因子选股的基本框架 ┊ 52
5.2 市值因子 ┊ 55
5.3 反转因子 ┊ 63
5.4 多因子选股策略的简单优化 ┊ 69

第 6 章 推进分析 ┊ 76

6.1 推进分析框架 ┊ 77
6.2 多层推进分析 ┊ 82
6.3 推进分析下的验证 ┊ 86

第 7 章 推进的择时策略 ┊ 89

7.1 均线趋势策略的推进分析 ┊ 90
7.2 均线反转策略的推进分析 ┊ 94
7.3 均线混合策略的推进分析 ┊ 96
7.4 自回归策略的推进分析 ┊ 99
7.5 自回归策略的多层推进分析 ┊ 102

第 8 章 推进的选股策略 ┊ 107

8.1 多因子选股策略的推进分析 ┊ 108
8.2 多因子选股策略的多层推进分析 ┊ 115

第 9 章 风险 ┊ 123

9.1 常用的风险度量 ┊ 124
9.2 其他风险度量 ┊ 129
9.3 风险和收益的结合 ┊ 134
9.4 止损 ┊ 138

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

第 10 章 仓位决策 ┆ 142

10.1 凯利公式 ┆ 143
10.2 实用的仓位决策方法 ┆ 152

第 11 章 仓位优化的择时策略 ┆ 155

11.1 仓位优化的均线趋势策略 ┆ 156
11.2 仓位优化的自回归策略 ┆ 169

第 12 章 投资组合决策 ┆ 181

12.1 最优投资组合理论 ┆ 182
12.2 实用的投资组合优化方法 ┆ 187

第 13 章 优化的股票配置策略 ┆ 193

13.1 多因子风险模型 ┆ 194
13.2 投资组合优化的多因子策略 ┆ 196

第 14 章 交易成本 ┆ 203

14.1 交易成本估计 ┆ 204
14.2 考虑交易成本的择时策略 ┆ 207
14.3 考虑交易成本的股票配置策略 ┆ 212

第 15 章 策略评价 ┆ 215

15.1 策略评价体系 ┆ 216
15.2 策略评价报告 ┆ 218

第 16 章 结语 ┆ 223

16.1 内容总结 ┆ 224
16.2 研发流程的局限与应对 ┆ 227

参考文献 ┆ 231

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

第1章

引　言

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

1.1 量化交易策略简述

本书的主要目的是为量化交易策略的研发提供一个可以参考的流程框架。但是直到目前为止，对于量化交易的界定，仍然存在着比较大的分歧。因此，为量化交易策略给出定义，虽然是全书开头一个不得不做的工作，却也是一件会导致争议的事情。为了尽量减小一个明确定义受到质疑的可能性，本书采用了相对而言较为稳妥的处理办法，即贴近字面意思进行解释。不过，出于一个研究人员的偏执，作者同时对策略的研发过程进行了着重强调，因此有了如下的定义：**量化交易策略，就是采用数量化手段构建而成并进行决策的交易策略。**

具体解释起来，该定义包括两层含义。

首先，在构建交易策略的过程中，数量化的手段应该占主要成分。这里的数量化手段，包括对整个交易流程和交易目标的数量刻画、数学模型的构建、对量化目标的最优化、对策略结果的数量化评价等方式方法。但是在这一部分中仍然会有定性的或者人为主观的成分存在，毕竟策略的研发是一个人为操作的过程。

其次，交易策略在构造完毕、用来进行交易决策时，必须具有明确的数量化规则，完全不存在主观判断的成分。这一特性使得整个策略可以在完全量化的设置下进行历史数据下的回溯测试，以及准确无误地指导交易操作，这些都是量化交易策略比较重要的特征所在。

同时满足以上两个方面含义的限定，则能够被称为量化交易策略。

在这样的界定下，量化交易策略既可以借助程序化的方式完成下单，也可以通过人工来执行。实际上出于成本和可控性等方面的考虑，一些交易频率较低的量化交易策略有可能更倾向于采用人工下单的方式来完成。在作者看来，策略的执行手段并不是量化交易策略的核心特征。

这种关注于数量化而非程序化的定义，也使得整个量化交易策略的历史比许多人认知中的要更长一些。因为就实际情况而言，技术分析中的技术指标，在适当的情况下是可以形成量化交易策略的。虽然技术分析中的图表分析手段，如“双头”“头肩”等图形形态的分类，相对而言太过主观而且很难量化[⊖]，但是技术

⊖ 罗闻全（Andrew Lo）等人通过对价格序列进行平滑的方式，完全量化地给出了这些价格形态的判别方法，再基于对当前价格的形态判断，可以形成相应的量化交易策略。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

指标这一门类侧重于价格和成交量的定量分析，通过公式化的计算得到一些用来参考的量化指标，进而指导交易，具有数量化的特征。只不过在这些量化指标的使用上，交易者往往又归于主观，进一步造成了对技术指标是否是量化交易策略的争论。

例如当某名交易员的交易策略是“移动平均线看起来很好时买入，看起来不好时卖出”，那么就完全有悖于上文给出的量化交易策略的定义。首先，该交易策略在表述上较为模糊，不是一个具有明确数量化规则的决策手段，因此交易员需要在交易过程中通过主观判断来完成买卖行为。其次，正是由于缺乏明确的数量化决策规则，交易员在形成这样的交易规则时很难定量化地描述整个交易策略和交易过程，也就难以使用最优化之类的数量方法。在多数情况下，交易员可能更倚重于复盘等人工形式来完成这一类交易策略的构建。

但是当交易员基于一些定量的规则来使用技术指标进行交易时，这些交易策略就可能会符合量化交易策略的特征。例如，把上面的策略改换为“价格线从下向上穿过移动平均线时买入，从上向下穿过移动平均线时卖出”，那么策略就既可以通过量化手段完成构建，又具有明确的数量化交易规则了。其他典型代表还包括大部分的技术指标，如理查德·唐奇安（Richard Donchian）所开发的通道规则，其以过去特定天数内的最高价和最低价为边界形成一个通道，当目前价格超出通道范围时，形成买卖决策。

理查德·丹尼斯（Richard Dennis）的“海龟交易法则”是一个非常著名的例子，因为丹尼斯招收交易员并传授该法则而为外界所熟知。这一交易策略正是在唐奇安通道指标的基础上构建而成的，除了通道突破的买卖规则外，海龟交易法则还包括仓位大小的选择、随时间的调整、止损等多个组成部分，更接近于一个构架完整的交易策略。

当然，就本书的定义而言，海龟交易法则是否算作标准的量化交易策略仍然有待商榷。首先，被披露的规则只是交易决策部分，具体的构建过程我们无从得知，因此也难以判断。更重要的是，其执行过程中存在人为主观的成分，这也直接导致了交易学员在使用同样规则的情况下获得了不同的交易结果。但这并不妨碍大量的从业者将海龟交易法则作为一个量化交易策略的范本来进行研究和使

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

用，作者本人也非常认同其在量化研究中的指导地位。

业内较为认同的量化交易策略的开端，可能是哈里·马科维茨（Harry Markowitz）的最优投资组合理论这一学术性创新[1]。在1952的论文当中，马科维茨开创性地引入了均值和方差这两个统计学上的概念，用来定量地描述投资者在投资组合上所获得的收益和承担的风险。基于投资组合中资产的收益情况和相互之间的协方差矩阵，采用最优化的方法，就可以得到投资组合的最优配比方式，从而帮助投资者在风险一定的情况下获取最大收益，或者在收益固定的情况下使得风险最小。整个过程通过明晰的数量模型进行了表述，优化结果也能定量地指导投资组合的构建行为，是比较契合本书关于量化交易策略的定义的。

在此之后，经济学家和金融学家开始越来越多地通过定量的数学模型来研究金融市场和投资交易。威廉·夏普（William Sharpe）等人于1964年前后在马科维茨的工作基础上，发展出了资本资产定价模型[2]。该模型将股票在无风险收益之上的超额收益分解为两个部分，即市场部分和残余部分，股票的风险也相应地分为两个部分，对应起来分别是系统风险和非系统风险。模型证明了在资本市场完全有效等前提假设下，残余部分的期望值为零，也就是说，非系统性风险是没有风险补偿的，股票的超额预期收益仅与其承担的系统风险大小有关。

之后，斯蒂芬·罗斯（Stephen Ross）从另外一些假设条件出发，得出了与资本资产定价模型在一定程度上具有相似性的套利定价理论。该理论同样认为股票的超额预期收益仅与其所承担的风险有关，只不过除了最主要的市场系统风险外，理论模型还能够包含其他一些存在风险补偿的风险因子。

在资产收益源于对风险的承担等思想的推动下，尤金·法玛（Eugene Fama）提出了著名的“有效市场假说”[3]，即在一个有效的证券市场中，价格完全反映了所有可以获得的信息，再结合资本资产定价模型的主要结论，认为所谓的消极型管理，即仅持有市场组合和无风险资产，才是真正明智的选择。随着这种思想

[1] 马科维茨因此而获得了1990年的诺贝尔经济学奖。

[2] 夏普因此而获得了1990年的诺贝尔经济学奖。

[3] 法玛获得了2013年的诺贝尔经济学奖。“有效市场假说”是其研究工作中最为核心的成果，而三因子模型也是其获得诺贝尔经济学奖的一个重要原因。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

在学术界的逐渐主流化，学术研究中对主动型交易的关注开始减少。

然而有趣的是，罗斯的套利定价理论和法玛的另一项著名研究却间接地发展出了一些行之有效的量化交易策略。在1992年的一篇论文中，法玛基于套利定价理论的形式，发现股票的两个当前特征（市值和账面市值比），可以有效地预测未来的股票收益。简单来说，就是市值越小的股票或者账面市值比越大的股票，在统计意义上会产生更高的未来收益。在次年的论文中，法玛将这两个特征构造为两个风险因子，再结合市场因子，从风险补偿的角度对实际数据进行了研究和解释，这项工作一般被称为“三因子模型”。

暂且不论市值和账面市值比这两个股票特征的风险化解释，1992年的论文中它们所表现出的显著预测能力，已经让逐利行为关注于此了。克里夫·阿斯内斯（Clifford Asness）是法玛在芝加哥大学指导的金融博士，其博士毕业论文在三因子模型的基础上加入了动量因子，以四因子模型的形式完成了一系列的实证分析。博士毕业后阿斯内斯进入高盛，成为一名量化交易分析员，随后在高盛组建了全球阿尔法基金，主要从事以量化为导向的交易工作，业绩不俗。1997年他离开高盛创办了自己的AQR资本管理公司，目前该公司是全球顶尖的对冲基金之一。虽然没有直接的证据证明阿斯内斯在工作中采用的是多因子模型基础上的股价预测技术，但是可以想见的是，市值、账面市值比、动量因子和因子模型应该与其量化交易策略存在一定的关联。阿斯内斯在一些访谈和学术论文中也时常谈到价值、动量/趋势、低风险、套息等相关概念，是为佐证。

由学术研究进入量化交易实业领域的一个更为极端的例子，应该是文艺复兴科技公司的詹姆斯·西蒙斯（James Simons），这也是中国读者较为熟悉的一个量化交易从业者。西蒙斯于1961年在加州大学伯克利分校取得数学博士学位，年仅23岁，并在30岁时就任纽约州立大学石溪分校数学学院院长。他在1978年离开学校创立了文艺复兴科技公司，该公司因为旗下的量化旗舰基金——大奖章基金傲人的业绩而闻名。关于西蒙斯所使用的量化交易策略，坊间一直有诸多猜测。许多人认为其所使用的应该是基于隐马尔科夫模型的量化交易策略，原因在于西蒙斯的早期合伙人伦纳德·鲍姆（Leonard Baum）是隐马尔科夫模型估计算法的创始人之一，同时文艺复兴科技公司招聘了大量的语音识别专家，隐马尔科

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

夫模型正是语音识别领域的一个重要技术工具。作者无法判断这一说法的真实性，不过不管怎样，从文艺复兴科技公司比较另类的人员构成来看，这应该是一个比较纯正的使用量化交易策略进行运作的对冲基金公司。

虽然大部分的量化对冲基金对其所使用的交易策略都或多或少地进行保密，但是仍然有一些量化交易策略在多年的应用后开始慢慢为外界所熟知，统计套利就是其中之一。这个策略的概念最早产生于摩根士丹利，当时的做法也被称为配对交易，实际上就是使用统计的方法选取一对历史价格走势相似的股票，当两只股票之间的价格差距变大、超出一定阈值之后，就分别做多和做空这两只股票，依靠该价格差在随后的时间里回归到正常水平来获取收益。由于这种量化交易策略既源于统计分析又存在等待价差回归的套利特性，因此被称为统计套利。而随着对这类交易策略的进一步深入研究，统计套利策略目前已经远远超出了配对交易的范畴，变得更加复杂和多样化。

大卫·肖（David Shaw）曾经是摩根士丹利这个统计套利交易组的成员之一，他于 1980 年在斯坦福大学获得计算机博士学位，随后留校进行学术研究。肖在 1986 年加入摩根士丹利后负责该组的技术部门，但是在两年之后，如同统计套利的首创者格里·班伯格（Gerry Bamberger）⊖一样由于政治斗争等原因从公司离职，并创立了自己的德劭基金公司。结合了肖的大规模并行计算研究背景和在摩根士丹利接触到的统计量化策略，德劭基金公司利用计算机量化模型作为主要的策略进行交易并取得了巨大的成功。

值得一提的是，肖在对冲基金领域扎根之后，仍然不忘科学研究，其成立的德劭研究公司致力于通过强大的计算机硬软件能力在分子动力学模拟等生化科研领域取得前沿性进展。这与阿斯内斯一直在金融杂志上发表学术论文的行为，相映成趣，当然肖的学术研究相对而言可能更为极客一些。

相比起统计套利，传统意义上的套利策略是一个更为人熟知、更经典的量化交易策略。实际上，现代金融框架的一部分都是基于“无套利”这样一个假设原则建立起来的，可见套利策略的深入人心与重要性。如果说统计套利的重点在

⊖ 关于谁首创了统计套利，存在若干种说法。综合来看，班伯格作为首创者的说法应该是较为可信的。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

于刻画和预测多个资产间的统计关系，那么传统套利可能就更注重各个资产的价值计算，以及策略执行时的交易成本估计和优化。只不过量化交易策略进化到现在，统计套利和传统的套利策略已经是互相渗透、互相融合的了。以计算机能力见长的德劭基金公司，对这两种套利策略应该都是有所涉及的。

说起套利，不得不提到长期资本管理公司。这家公司的阵容十分豪华，包括债券套利的先驱约翰·梅里韦瑟（John Meriwether）、两位诺贝尔奖获得者罗伯特·默顿（Robert Merton）和迈伦·斯科尔斯（Myron Scholes）、美联储副主席戴维·穆林斯（David Mullins）等诸多顶级从业者，主要从事的正是债券的量化套利交易，当然其中也会包含一些其他形式策略的成分。公司成立的前三年表现非常出色，但是在 1998 年俄罗斯国债违约之后，相关的连锁反应使得公司产生巨额亏损，在美联储的干预下被华尔街几家公司出资接管，形同倒闭。实际上，长期资本管理公司在俄罗斯债券上的损失本身并不大，但是许多大金融机构在亏损环境下必须保证足够的资本量，因此通过出售流动性较好的七大工业国债券等资产来降低风险、增加资本，全球主要债券价格在卖出压力下大幅下跌，波动巨大，这才导致了杠杆极大的长期资本管理公司在债券套利上产生巨额亏损。

这里谈及长期资本管理公司，决然不是为了说明套利策略的缺陷，或是用一个失败的量化交易特例来衬托其他量化交易公司的成功。实际上，套利策略是一个具有内在金融逻辑的优秀交易策略，只不过任何量化策略都难以完整地考虑到所有可能出现的情况，长期资本管理公司也正是被一个极端事件的连锁效应所击倒。一个量化交易的研究者和从业者，首先应该具有概率、统计的思维方式，应该理解任何有概率的事情都是可能发生的。对于交易这个行业而言，不到退出市场的那一刻，永远都存在着失败的可能，拥有再光鲜的历史业绩也无法摆脱这一事实，而且光鲜业绩本身可能也只是大样本下的一个幸运个体而已。希望读者能在阅读本书时始终保持这样的思维，不要盲目地被看似美好的回溯测试结果甚至实盘业绩所迷惑。

套利策略中一个非常重要的部分就是对交易成本进行判断，同时尽可能地降低交易成本，从而保证足够的套利空间。随着交易电子化的不断发展，以及美国证监会“另类交易系统规定”等法案的推出，交易策略的自动执行也在持续地发

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

展和演化。在这样急速更新的市场环境下，开始出现一些相应的量化技术，例如通过订单拆分来减小市场冲击成本的算法交易等。

更为人熟知的是高频交易这个概念，交易的执行者通过计算机下单、将策略执行部件放置于距离交易所主机更近的地理位置上、交易订单直通交易所等手段，从时间延迟等层面降低了价格变动带来的交易成本，从而增加套利空间。当然，在作者写作本书时，高频交易技术已经进入微秒级别的领域，高频交易也被更多地使用在了做市商策略等其他策略之上。

在开头处已经提到，本书的主要目的是为量化交易策略研发提供一个可供参考的流程框架。而高频交易等策略方法在很大程度上更侧重于执行层面的技术手段，对比传统意义的量化交易策略具有一定的独立性，因此本书将不涉及高频交易、算法交易等量化交易策略的内容，对套利的讨论也局限于非常浅显的层面。不过，作者仍然认为中低频的量化交易策略研发框架对于高频交易同样存在一定的借鉴意义，原因在于，开发交易策略始终是离不开对收益和风险的认知的。由于高频交易的内容不会出现，因此从此处开始，书中但凡提到量化交易策略，如没有具体说明，均特指中低频量化交易策略，敬请注意。

1.2 量化交易策略的优缺点

查尔斯·达尔文（Charles Darwin）在 1859 年出版的《物种起源》一书中系统性地阐述了他的进化论思想，认为生物是进化的，而自然选择是生物进化的动力。如果一个有利的变异更能适应当前的环境，那么变异的个体就更容易生存下来并繁衍后代，保存自身的性状，从而不断地进化。作者一直都很倾向于将这一理论类比到某种技术发展或者社会实践上，一种事物的产生并兴起，往往是因为它更好地适应了环境。好像量化交易策略一样，因为学术理论的不断产生、交易环境的不断更新、计算机技术的不断进步而持续发展。而由于量化交易策略在这种大环境下表现出了种种优点，以及对比其他交易策略的一些优势，从而呈现出了欣欣向荣之势。

鉴于量化交易策略自身的特质，其最为明显的特点就是较强的可度量性。作

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

为采用数量化手段构建而成并进行决策的交易策略，其在构建过程以及决策过程中，都是可以被精确度量的。对比而言，主观化、定性化的交易策略虽然在复盘等过程中也能够得到数量化的交易结果，但是由于缺乏整体性的精确度量，因此局部的定量化结果往往波动较大，在很大程度上不具备稳定的刻画能力。

这一特性带来了量化交易策略的另一个优点，即可验证性。由于未来数据的不可得，因此不论是定量的还是定性的，实际上交易员在判断一个交易策略时，很大程度上都依赖于策略在历史数据上的表现结果。然而，主观的交易策略在复盘或者其他测试过程中含有人为判断的成分，因此在缺乏过往实盘操作时，历史验证就成为一个不可精确重复、缺乏稳定性和说服力的部分，这对于此类交易策略而言是一个比较严重的缺陷。

量化交易策略就不存在这个问题，基于本书给出的量化交易策略的定义，只要准备好策略所需的完整数据以及资产的历史价格，我们就可以将交易的过程在历史数据上模拟复现一遍，这一过程被称为回溯测试，简称回测。在策略构建和策略表述数量化的基础上，多次重复历史回溯测试均能够得到一致的结果。如果测试结果是正向的，那就说明该量化交易策略在历史检验上具有盈利能力。

上述两个特性使得量化交易策略具有较强的客观性。由于构建交易策略的过程中数量化手段占主导地位，交易决策更是具有明确的数量化规则，因此量化交易策略可以在很大程度上规避策略开发者的主观臆断，在构建的过程中始终获得客观地对待。这里举一个反例，作者曾经在构造一个量化交易策略时，由于急于获得最终的结论，因此对策略粗略运行得出的待加工结果进行了人工观察。在没有进一步编写程序进行验证的情况下，得到了该策略可以获利的结论。但是在测试完整策略的时候，发现预期情况和实际不符，重新检查后才发现偏差来自这一点点的人工行为。量化交易策略的研发已然如此，作者有理由相信在主观策略的研究过程中，这种情况更为常见。实际上，在开发人员对策略有效性有期待的情况下，即使面对一个非常简单的问题，人的主观判断也是有可能会产生偏误的。

同时，明确的数量化交易决策规则这一特点，也使得交易员在执行量化交易策略时，有希望完全排除情绪对于整个交易过程的干扰和影响。基于数量化规则，量化交易策略本身具备相当的独立性，不需要交易员的主观判断也可以完整

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

地指导整个交易流程。而主观的交易策略由于在交易操作中需要持续的人为控制和判断，因此不可避免地存在一定的偏向。行为经济学里研究了很多相关的问题，包括损失厌恶偏向、过度自信偏向、参考点偏向等，都是一些人类固有的行为偏向。当这些客观存在的偏向影响到交易策略的实际执行时，交易结果的预期就会从最优点处发生偏移。虽然量化交易策略不可能帮助我们完全地规避这些问题，但是数量化的框架和规则确实可以尽量减少这些不安定因素所带来的损害。

上面所说的排除情绪化操作，实际上也包含了交易员常常谈到的一致性，也就是保证交易策略在执行过程中，使用相同的交易规则，其中包括买入点、卖出点、交易仓位大小的确定等。如果交易员在执行量化交易策略时不主动加入人为判断，量化的交易规则可以帮助其很轻松地完成这个任务。更为重要的是，量化交易策略可以做到历史验证过程和实际交易行为的一致性，因为不论是真实交易决策，还是历史回溯测试，所参照的交易规则都是由数量化表达精确定义了的。这种整体上的一致性，是绝大部分主观交易策略都无法保证的。

量化交易策略的数量化特点及其带来的一致性，使得这种策略具备了比较好的可移植性。除非使用特定的量化因素，一般而言，量化交易策略是比较容易在一个市场或者资产上证实有效后，移植到其他的市场或资产上使用的。例如，当策略使用的量化数据仅限于价格时，所有存在报价的市场都能够使用历史数据来验证这种量化交易策略。而当使用的量化数据包括价格和成交量时，除了外汇市场等少数情况外，量化交易策略也是能够移植到其他大部分的场内交易市场上的。量化交易策略所使用的数据的可得性越强，其移植能力也就越强。

同时，由于量化交易策略可以很方便地得到数量化的验证结果，收益、风险等策略特征都能够通过数据形式得以呈现。因此，在量化交易策略被应用到多个市场或者资产上时，策略开发人员可以通过数量化结果这样的直观形式来进行横向对比，从而有针对性地选取适合某一特定量化交易策略的市场，或者在多个市场之间进行配比。而由于量化交易策略的执行过程中不需要交易员的主观判断，因此相比于主观化的交易策略，其有能力同时覆盖大量的市场和资产，这对于投资组合的分散化来说具有很好的辅助作用。在现实中，大部分的量化基金都持有数量较多的资产和资产种类来组成投资组合，这正是利用了量化交易策略的特性

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

从而更方便地分散风险。

另一方面，对于量化交易策略而言，这种多个市场间的策略移植所需要的成本很小，在一些情况下甚至不需要变动原先的基础工具，因此降低了研发成本，同时节省了宝贵的研发时间。而量化交易策略不仅可以在策略研发的过程中帮助节省时间、降低成本，由于其具有非常明确的数量化规则用以指导交易，因此策略在实际使用的过程中可以脱离人为判断，执行速度更快，运作效率得到了提高。不论是否采用程序化的执行手段，量化交易策略都能够在实际交易中减少人的负担，也就减少了许多重复性的劳动。对于一个交易策略的控制者而言，就能够把更多的精力放在最为核心的策略创新上。

而在策略创新上，量化交易策略本身也具有一定的优势。在科学不断进步的今天，越来越多的先进技术被创造出来并应用到各种情境之下。通过结合多个不同学科的知识和相应的数学模型，量化交易策略更有可能发现一些隐藏较深的复杂数据规律，而这些规律往往不太容易被主观交易者察觉得到。时至今日，神经网络、支持向量机、隐马尔科夫模型等机器学习模型已经开始被量化交易从业人员频繁提起，虽然作者并不能确定这些技术手段在量化交易策略方面有多大程度的应用，但是可以肯定的是，交叉学科知识进入量化交易策略研发已经形成一种趋势。而量化交易策略本身也在融合这些科学技术的同时保持极快的发展速度，以至于有些从业人员开始用“第几代量化交易策略”之类的说法来划分各种策略方法。尽管作者并不太赞同这种过分注重技术手段的行业观点，但是量化交易策略当前的迅速发展是不可否认的。当然，主观的定性交易策略也是在不断进步当中的，只不过在目前这种科技发展、学科交叉的环境下，量化交易策略发展和进步的速度应该是更快的。

实际上就整个行业的状况而言，量化交易策略还有一个不太被人提及的优点或者说特点，那就是能够在一定程度上帮助公司减小对所谓的明星交易员的依赖。由于在交易流程当中，做决策的不再是具体的交易员，而是已经具有明确规则的量化交易策略，因此在量化交易策略研发完毕甚至只是核心理念构建完成之后，公司对于量化交易策略的研发者的依赖程度就会迅速下降，这一方面可以帮助公司在人员配置上减少支出，更重要的是提高了公司对于交易整体的控制程

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

度。整个公司系统的核心优势在于量化交易策略本身，因此不会由于人员变动等问题而非常迅速地消失，有利于公司在量化交易方面保持长期稳定。

不过这个特质在稳定公司组织架构的同时，实际上也损害了量化交易策略研发者的利益。量化交易策略的研究人员在策略研发完成之后丧失了一定的价值，因此在薪资等方面也不如主观交易的明星交易员一般具有讨价还价的底气。这也导致一部分的量化交易从业者较为消极地对待工作，甚至选择保留核心交易策略，进而隐性地损害了公司的利益。好在很多情况下，量化交易策略即使在研发完成之后，仍然需要不断地改进，从而适应市场环境的不断变化，因此对量化交易策略研发人员的持续需求仍然是存在的。当然，这个问题要妥善解决，仍然在于公司对此的激励政策。

量化交易策略由于其明确而数量化的交易规则，不仅会造成上面提到的这一缺点，更重要的是，这些明确的量化规则是非常容易被复制的。量化交易策略和许多科技创新一样，具有研发困难、复制简单的特点。即使在最终交易规则保密程度良好的情况下，仅仅只是泄露出一些策略理念出来，相比于主观交易策略，量化交易策略也有更大的可能会被逆向工程破解从而外泄。这种特性，一方面增加了量化交易策略使用的成本；另外一方面也加大了量化交易策略在交易之外的风险。一些量化交易的从业者在交流过程中非常小心谨慎，公司在管理量化交易时执行诸多规章条例、对员工加以合同上的严格限制甚至最后走到对簿公堂的境地，都是由量化交易策略这样的特性引起的。

量化交易策略还有一个非常显著的缺点，那就是量化这个特性本身。诚然这一特性为量化交易策略带来了多种优势，但是由于这一特点，使得量化交易策略在对待无法量化的因素时，只能采取忽略、舍弃这样较为无奈的处理办法。因此，量化交易策略损失掉了很多其实有可能带来盈利的信息，也使得策略在处理信息时所覆盖到的范围变得狭窄。当然，随着科学技术的发展，一些之前无法被量化的因素开始进入量化交易策略的研究范围之内，如针对网络微观信息刻画出的投资者情绪等。但是，即使科技手段会让量化交易策略能够处理的信息范围更广、更深，相对于主观交易策略而言，这样的缺陷始终是量化交易策略无法彻底摆脱的。这种来源于自身特性的缺点，目前来看只能改善而无法根治。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

同时，由于量化交易策略在构建过程中采用的是数量化的方法，因此除了套利策略这种更注重模型和逻辑的策略之外，大部分基于统计、优化等手段的量化交易策略，均需要一定数量的数据样本进行研究，而相应的数据都是随着时间的推移逐渐产生的，所以当量化交易策略的构造形式没有本质上的改变时，从数据中抽取的数量化特征也只会随着时间的延长逐渐变化，策略所形成的交易也只能缓慢地变化。当市场情形发生重大转变时，这种缓慢变化的特性会导致量化交易策略无法适应转折期的市场，在短时间内造成较大的损失。

而对于套利策略来说，当其缺乏主观判断控制的情况下，结构更为稳定，转向也更为缓慢。如果出现某些小概率的突发事件，有可能就会产生如长期资本管理公司那样极端负面的结果。相比较而言，一部分定性的交易策略由于主要基于逻辑上的思路来进行交易，因此当市场情形发生改变时，能够基于主观逻辑迅速进行本质上的策略调整。转向缓慢这一特点也是量化交易策略一个比较难以改善的缺陷。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

第2章

量化交易策略的研发流程

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

2.1 量化交易策略的基本研发流程

本书将在后面的章节中，通过流程化的说明和一些中国市场上的实际例子来介绍量化交易策略的具体研发框架。通过整本书的内容，读者将遵循一定的逻辑来具体了解量化交易策略是怎样构成的以及是如何被研发出来的。

在本章中，作者在精炼量化交易策略基本研发结构的基础上，通过介绍有限的几个组成部分，对一个较为完整的量化交易策略进行了初步说明。希望读者通过阅读体量不大的文字内容之后，对量化交易策略的研发形成一个大致的概念和逻辑上的了解。图 2-1 给出了一个最为基本的量化交易策略研发流程，这是一个较为松散但同时相对典型的研发流程，对于本节中的基础性介绍而言较为合适。

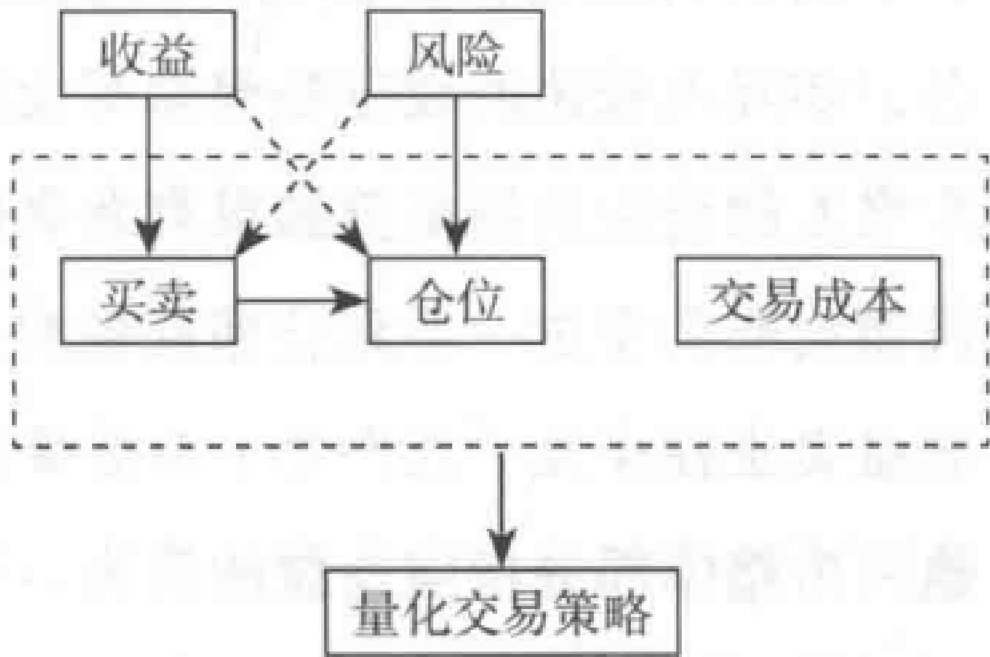


图 2-1 量化交易策略的基本研发流程

在整个量化交易策略的研发流程中，买和卖是最为基本的量化交易策略组成部分，而这个部分的设定主要与收益情况相关。这里所说的相关，具体分为两种不同的情形，一种是总体的关联性，即基于买点、卖点的选择，买卖策略应该得到一个正的整体收益。另一种则来自对交易资产未来收益的判断或者说预测，即判断交易资产的未来收益为正时，就买入资产；判断交易资产的未来收益为负时，则卖出或卖空资产。

实际操作中，这两种关联关系的情况可能更为复杂一些，读者可以通过第 4 章的内容进行具体的理解。有的时候，买和卖的具体操作也可能受到风险方面设置的影响，例如为了限制单次交易的最大损失而采取止损之类的操作时，用于清仓的买卖设置就会相应地变动，这也是作者将风险和买卖用虚线相连的原因。不过在大部分情况下，买卖这一最为基本的组成部分还是与收益的关系更大，研究者也应该在研发这个组成部分时，着重考虑收益情况的具体影响。

对量化交易策略风险的控制可能会影响到量化交易策略中的买卖设置，但是在更普遍的情况下，风险这一因素主要影响的是交易仓位的设置。当然，前提条件还是需要买卖策略的总体收益为正，在这样的条件下再结合仓位的设置，才能

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

够在合适的风险水平下取得达到要求的收益。

通过对交易资产具体仓位的调整，交易者可以比较直接地控制单次交易以及整个交易策略的风险水平。例如在满仓交易的情况下，定量地判断了当前交易的风险之后，觉得风险过大无法承受，那么最为直接的处理方法就是在满仓的基础上相应地降低仓位的大小。在仓位降低之后，对于整体资金而言风险也就随之降低了。由于仓位本身具有量化、直观的特性，因此当交易者希望将风险处理到一个特定的水平时，调整仓位是一个比较方便的手段。

需要说明的是，前面已经提到了买和卖是量化交易策略最为基本的组成部分，实际上仓位的设定是根据买卖决策和风险两个因素共同形成的，不建立在买卖之上的仓位选择是空洞没有意义的。此外还有一个更为极端的情况，仓位的正确设定有助于进一步优化策略的整体收益，第 10 章 10.1 节所要介绍的凯利公式的意义正在于此。在图 2-1 中由买卖到仓位的箭头，实际上可以看作收益、买卖这一个整体部分指向仓位的箭头。不过在实际使用中，凯利公式所导出的仓位设定往往过于偏激，超过正常风险控制下的最高仓位值，因此仓位仍然与风险的关系更为紧密。

在图 2-1 这个较为松散的量化交易策略研发流程中，交易成本是和买卖以及仓位具有同等地位的组成部分。在实际操作中，就是首先基于对收益和风险的判断得出合适的买卖和仓位选择，然后在买卖和仓位共同组成的量化交易策略当中考虑交易成本，也就是在建立仓位和退出仓位等操作中扣除所需要承担的交易成本。随后再次判断该量化交易策略所代表的收益和风险情况，若这两个因素仍然在接受范围之内，则确认这是一个可行的量化交易策略。虽然最后用来执行的组成部分只有买卖和仓位，但是交易成本作为对量化交易策略的一个实际化修正，也是策略研发流程中一个不可或缺的组成部分。

上面提到的对量化交易策略收益和风险情况的判断，实际上是一个综合性的评价问题。一个最为重要的参考依据应该是策略在整个交易过程中的净值走势，通过对策略净值走势的分析，就可以建立起该量化交易策略运行情况的全面判断。但是净值走势本身由于细节过多，因此无法简单地用来进行策略之间的横向对比。这时就需要精炼净值走势中所包含的信息，选取合适的部分形成量化的评

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

价指标，从而进行量化交易策略的进一步判定。

就作者看来，评判一个策略的标准中最重要的仍然是策略在整个交易过程的收益情况，一个负收益的量化交易策略根本无须考虑其风险即可排除。而当收益率为正时，再结合风险的度量进行具体的取舍，就可以直观地给出量化交易策略是否合格的评判标准了。净值的回撤水平是作者在研判交易策略净值时比较偏好的一个风险度量指标，在后面的案例分析中我们也会重点查看回撤的结果。

与此同时，一些量化交易策略在进行收益和风险情况的判断时，仅仅针对策略自身的净值走势进行研究是不够的，给出一个合理的基准来进行对比往往是更为有效的判别方法。例如，后面的案例中会涉及量化选股策略，当交易选择仅限为对具体的股票进行持仓，而不考虑空仓或者卖空时，选取一个特定的基准进行对比就会是一个更为有效的判别方法。这主要是由于量化选股策略的仓位始终为多头，因此不论如何配置，策略所持仓位都含有资本资产定价模型中所提到的市场成分。而选股策略本身的意义在于选择更好的股票，不在于获取市场收益，因此将市场走势作为对比或者在策略收益中剔除掉市场成分就是一个更合理的做法。

第 5 章和第 8 章中的因子选股策略就使用了沪深 300 指数作为基准进行了比较，同时简单地用策略收益率减去指数收益率形成新的收益率序列，进而研究该收益率序列所形成的净值走势情况。当然，由于做空也需要保证金的投入，因此这样简单的减法处理虽然直观有效，但同时也欠缺一定的实际操作意义。一个更加贴合实际的做法是以市场中性组合的方式进行策略构建，进而研究中性策略的净值走势，有兴趣的读者可以自行进行研究。

上述所有的操作，都需要建立在对历史数据的分析之上，也就是前面章节提到过的回溯测试。所谓回溯，是将交易的过程在历史数据上复现一遍，这里面包含一个假设，即历史数据在量化交易策略中展现出的样本特征在未来的交易中依然存在，否则回溯测试就失去了意义。关于这一假设的分析其实在诸多技术分析著作中均有涉及，一般被称为“历史会重演”，这里不再继续展开。

不同于传统技术分析的是，量化交易策略的研发过程更加深入具体，在涉及

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

策略的参数设定、模型设置等具体问题时，需要采用数量化的方法进行解决。例如，如何设置买点和卖点可以使得相应的总体收益最大、如何设置模型参数使得收益率预测最为准确等，都是很典型的最优化数学问题。那么找到合适的模型、最优化技术和算法并加以应用，就能够确定量化交易策略的最终形式，用以进行实际交易。

2.2 量化交易策略研发流程的进一步论述

2.2.1 量化选股策略的研发流程

图 2-1 中所展示的是一个较为松散的一般性框架，用来总领性地说明量化交易策略的基本研发流程。在具体的策略研发过程中，这个框架经常会因为具体研发设置和策略设置的不同而产生变化。例如，当量化交易策略的主要作用不是在时间轴上选择具体的买卖时点，而是在同一个时间点上对多个资产进行选择和配置时，图 2-1 中的一些说明就显得有些含混不清。

在本书中，作者将在后面介绍的选股策略就是这一类策略中最为常见的形式，因此这里在整体框架不变动的情况下，针对图 2-1 进行了文字上的调整，用以说明量化选股策略的运行框架与研发流程。当然，使用选股策略的框架体系来处理多个资产甚至多个策略的挑选、配置也是可以的，在不复杂的情况下读者只需要稍做联想即可。

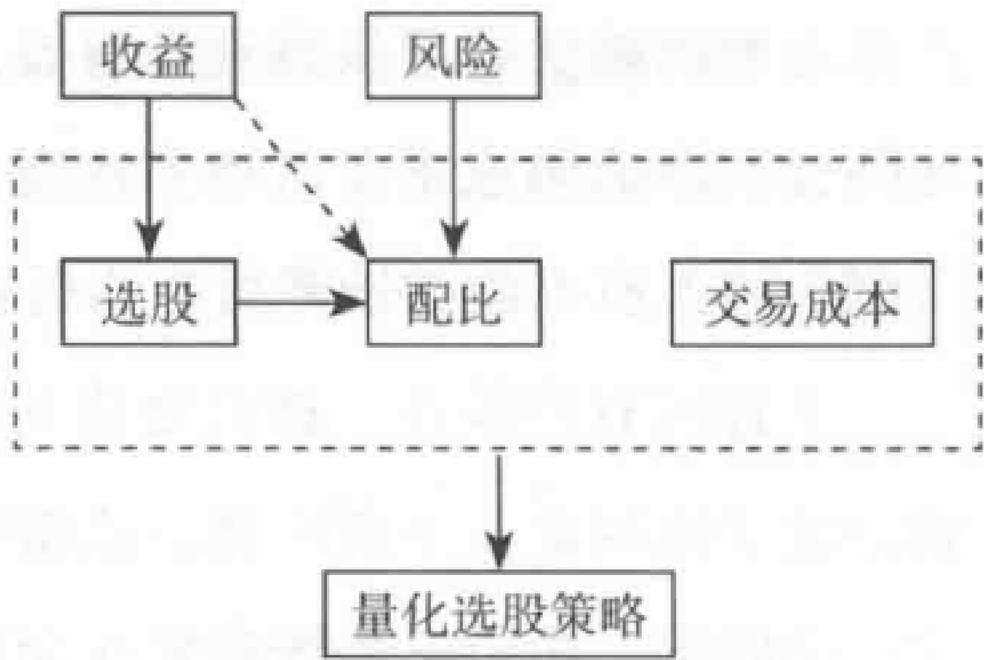


图 2-2 量化选股策略的研发流程

买卖和仓位虽然是更为通用的说法，但是更适合于描述择时策略，放在选股策略的研发框架中会显得比较突兀，图 2-2 将买卖换成了选股，仓位则换成了配比，这样更容易让读者领会该研发流程的含义。实际上，对于每一期的选股而言，如果选择了原先没有仓位的股票，那么对应的操作就是买入该股票，如果已经建仓的股票没有被选入这一期的股票池，那么对应的操作就是卖出该股票。而配比则是在

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

买卖的基础上，通过仓位大小的变化来实现具体配置。因此，选股和配比实际上可以算作买卖和仓位选择的特殊情况，只是这种说法更为贴合量化选股策略本身。

风险在量化选股策略研发流程中的具体含义则略有不同。由于选股策略的仓位操作涉及多个股票之间的配比问题，因此这里的风险不仅包括单只股票的风险，也涉及多只股票之间的风险程度，后一种风险一般采用股票收益之间的相关性来进行描述。例如在一般性的最优投资组合理论当中，经常使用协方差矩阵来刻画整个资产组合的风险水平。

虽然从实际情况来看，相关性这一度量方式与风险的直观感受之间有一定的差距，但是在多资产环境下，一般都将资产间的相关性视为风险的来源之一，这是一个偏学术的、约定俗成的做法。

就实际意义而言，我们可以这样理解。当我们知道了一个投资组合中各个资产间的协方差矩阵之后，就可以根据公式计算出整个投资组合的方差，这代表了投资组合这一整体的风险水平。其他情况也是类似的，当我们已知多资产的风险和资产间的相关性程度时，就可以估算出投资组合整体的风险。

由于这种风险刻画的方式主要针对多个资产的投资组合，因此在第 9 章论述风险时并不涉及，而在第 12 章讨论投资组合决策时进行有针对性的说明，并在第 13 章介绍实际的量化选股策略时给出具体的协方差矩阵估计方法。

2.2.2 量化交易策略研发流程的变化形式

上面的例子是针对选股策略进行的文字上的变动，实际上量化交易策略研发流程的变化更多来自各个研发组成部分不同的结合方式。而不同的结合方式，对应的是策略研发过程中不同的目标和需求。

例如图 2-1 所介绍的松散的研发流程，是在确定好买卖行为和仓位设定之后，再针对实际交易中所产生的交易成本进行二次测试。这样的做法虽然简便易行，但是忽视了交易成本本身对于收益的影响，以及更进一步对于买点和卖点的影响。

因此，在确定买卖设置的步骤中就考虑交易成本的影响，应该是一个更贴近

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

实际的研究框架。特别的，当某些情况下交易成本对整体业绩影响非常之大时，如交易频率较高的情况，这样的考量就显得很有必要了。

图 2-3 给出了相应的流程刻画，如图所示，在判断收益因素时，同时考虑交易成本对于收益的影响，从而优化出更为实际的买卖设置。再根据相应的风险控制，结合买卖点的选择，得出最后的仓位设置。在确定了买卖和仓位这两个部分之后，就获得了一个完整的量化交易策略。

图 2-4 给出了一个更紧凑、更贴合实际操作的量化交易策略研发流程。在该流程中，买卖和仓位的设置是同时作为参数进行优化的，优化的目标函数也进行了唯一化，即量化交易策略的风险调整后收益。而在确定需要优化的目标函数时，交易成本也如同上一个研发流程一样同时被考虑进去，从而保证买卖和仓位优化结果的准确性。

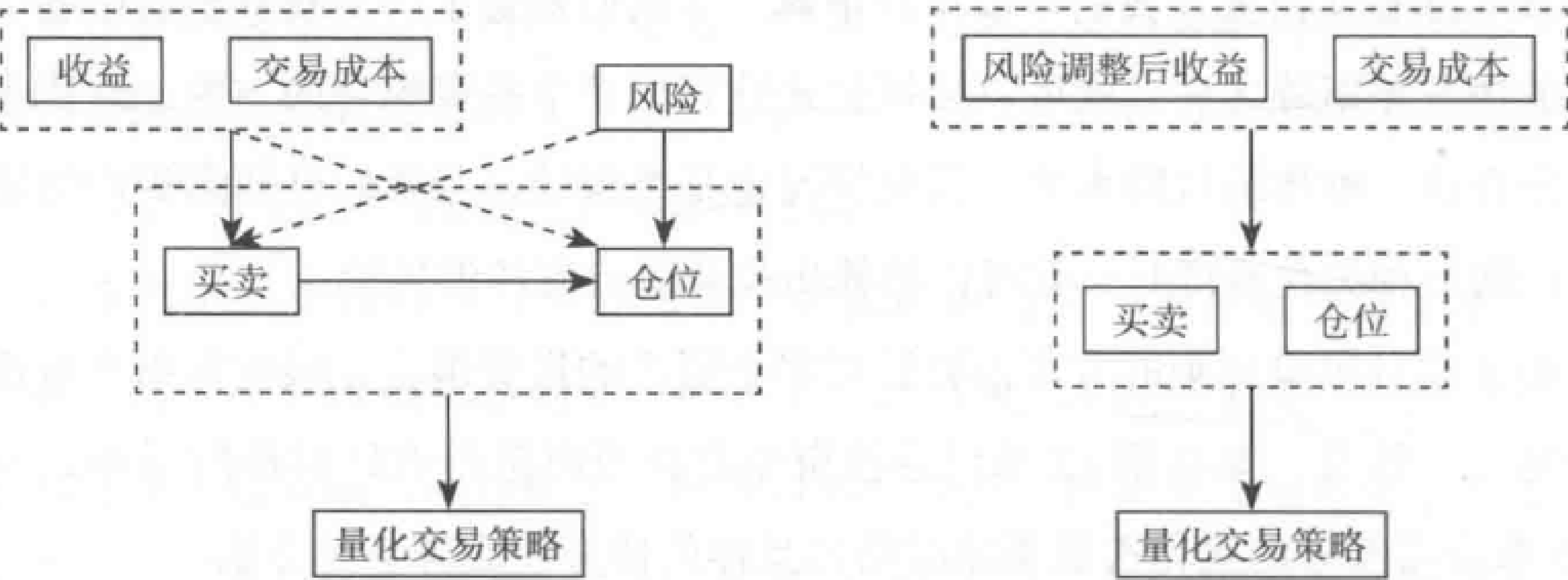


图 2-3 收益和成本共同决定买卖的研发流程　　图 2-4 买卖和仓位同时优化的研发流程

毫无疑问，相较于上面所涉及的研发流程，特别是图 2-1 中较为松散的研发流程，该量化交易策略研发流程的各个组成部分更为紧密，因此在优化过程中所产生的与实际操作的偏离也就越小，买卖和仓位设置的准确度也就更高。但是在实际工作中，如果想参照这一流程进行研发，那么就需要比较强的计算能力，数据量的大小也要达到一定要求，同时优化方法和目标函数的设定要能够同时覆盖买卖和仓位的所有参数，因此往往也只有极为简单的策略可以采用这样的流程框架进行研发。例如后面会详细介绍的双均线策略，实际上是可以通过这样的研发流程来完成的，有兴趣的读者可以自行研发。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

2.2.3 量化交易策略的整体工作流程

在实际的量化交易策略相关工作中，研发只是整个工作流程的一部分，还有两个组成部分需要着重强调。基于此，图 2-5 在图 2-1 所示的研发流程的基础上给出了一个更为完整的工作流程。如图 2-5 所示，需要增加的部分包括处于研发过程之前的数据准备工作以及处于研发过程之后的策略执行工作。这两项工作与前面所论述的研发流程具有很强的逻辑关联性与内在依赖性，三者结合起来形成的一个整体，基本上可以涵盖量化交易策略具体工作的绝大部分内容。

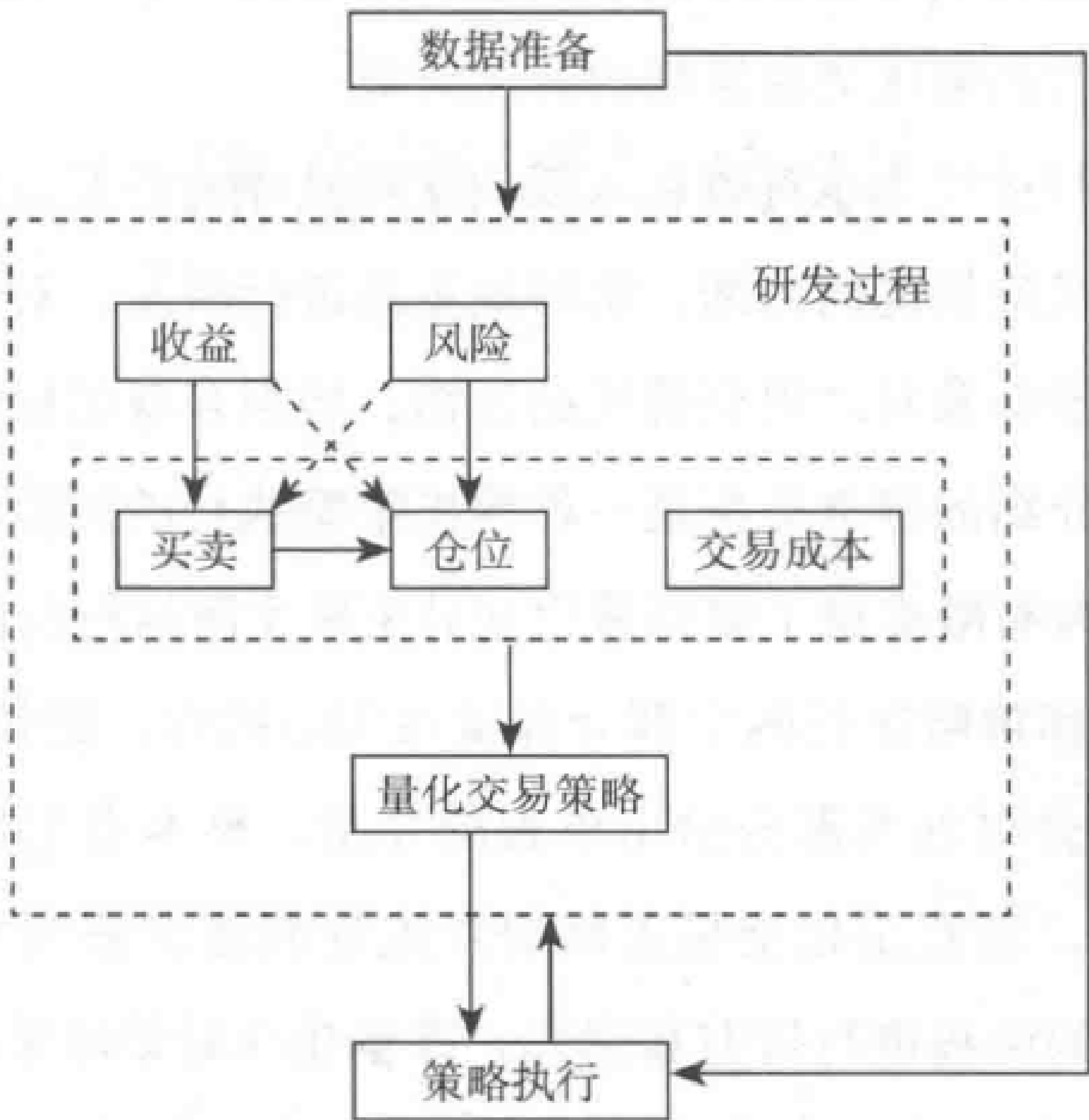


图 2-5 量化交易策略的整体工作流程

首先论述数据准备的工作，循着图 2-5 中的箭头可以看到，在量化交易策略的整体工作中，既要为研发过程准备相应的研究数据，也要为策略执行准备相应的实时数据。

在研究数据方面，由于寻找合适的量化交易策略需要不断重复研发流程，因此对于数据的要求更侧重于准确性和覆盖能力。同时，对数据的清洗和转换也是一项重点工作，在大部分的数据科学研究、包括量化交易策略的研发当中，数据特征的合理抽取对于整体效果提升的重要性有时甚至要高于精巧的模型，当然很多时候数据的转换和模型的构造是相互融合的，针对具体情况应当采取具体的分

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

析和处理。

而在策略执行数据方面，则更应该关注数据获取的及时性。至于数据的清洗和变换，只需要完全复制研发得到的量化交易策略下的数据准备工作即可。另外，为了保证数据的及时性，最终进行的数据清洗工作对时间消耗存在一定的要求。

然后，讨论策略执行的工作。策略执行，是在量化交易策略研发完成之后，最终产出实际效能的组成部分。执行时应该遵循尽量贴近研发完成的量化交易策略的原则，与量化交易策略所确定的买卖、仓位等设置尽可能保持一致，这样才能最真实地反映出前面量化交易策略的研发结果。

同时，策略执行的结果也可以用来反向支持具体的研发流程，通过对策略执行所得到的收益、风险情况的判断，实时地重新进行研发，对量化交易策略进行修改，从而使策略能够及时地得到现实的反馈，增强自身的稳健程度。值得一提的是，第 6 章将要介绍的推进分析是一种模拟策略执行的回溯测试技术，读者可以在运行推进分析时有限度地了解到量化交易策略实际执行时的种种状态。

由于数据准备和策略执行两个部分都重在实际操作，比较难以呈现在纸上，因此后面的内容不会对这两部分给出单独的介绍，整本书主要集中讨论量化交易策略的研发流程。而在量化交易策略研发流程的诸多框架类型当中，则选取作者认为最为基本的流程进行研究和阐述，即量化择时策略采用图 2-1 的研发流程，量化选股策略采用图 2-2 的研发流程，希望读者能够根据最为基本的框架举一反三，并在实践中慢慢掌握如何根据具体的研发环境选择最为合适的流程框架。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

第3章

注意事项与应对

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

3.1 未来信息的规避

在量化交易策略的研发过程中，存在许多需要注意的事项。如果忽视了这些问题，轻则造成量化交易策略的优化结果与实际要求存在偏差，重则直接导致最后的量化交易策略没有任何实际用处，甚至造成真实的损失。

在这些事项中，首先需要注意的是量化交易策略研发中是否存在未来信息的问题，如果对历史数据进行建模和优化的过程中包含了未来信息，那么应该重新构建研发模型，规避掉所有的未来信息，保证量化交易策略与实际相符。

所谓未来信息问题，就是指基于历史数据进行回溯测试的过程中，使用交易完成之后的信息来决定交易本身。为了更为直观地说明何为未来信息，这里给出了一个用来解释的示意图，如图 3-1 所示。

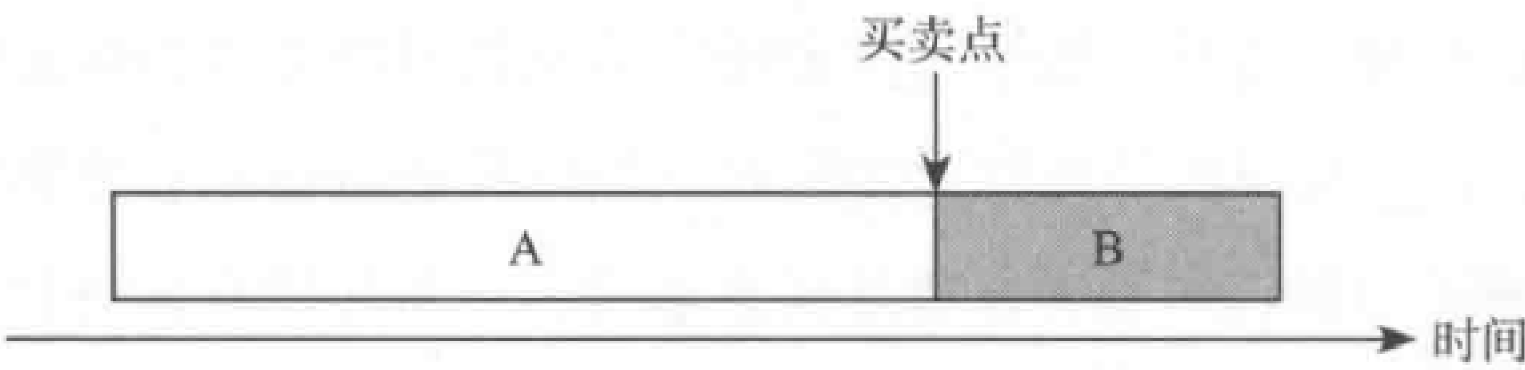


图 3-1　未来信息示意图

在图 3-1 中，回溯测试能够使用的样本数据包括 A 阶段的数据和 B 阶段的数据，假设模型判断出一个买卖点在 A 和 B 之间，那么 B 阶段所有的信息对于这一个买卖点而言都是发生在未来的信息，因此在判断交易的时刻，B 阶段的所有信息都是无法得知的。

在量化交易策略正常执行的过程中，买卖、仓位等策略组成部分的判断都是基于之前所有可以获取的数据得到的，也就不存在未来信息的问题。但是在进行回溯测试的时候，由于我们已经得到了超越买卖点的数据信息，也就是图 3-1 中 B 阶段的信息，因此就可能发生将 B 阶段的未来信息用以判断买卖点这样一种不符合实际交易逻辑的情况。

在国内的一些资料当中，描述这种问题更常见的名称是未来函数，作者猜测这样的习惯可能源于国内交易者对于技术指标的深入研究。在编写技术指标函数的时候，往往会由于疏忽等问题出现这种使用不应该出现的未来信息的情况。而

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

在国外的一些文献和资料当中，比较常见的说法还包括前视偏差。相较于未来函数而言，作者认为这一名称的覆盖面更加广泛一些，因为在一个较为完整的量化交易策略研发过程当中，函数只是其中的一部分。而前视偏差则比较形象地说明了这个问题，站在图 3-1 中的买卖点上，B 阶段的信息都是时间前进后才能得到的，这样所造成的偏差就会让未来的信息进入交易决策本身。

使用 K 线图当前柱的最高价、最低价判断买卖点是一个比较常见的使用未来信息的情况。举例而言，当一个建仓策略为“若今日的最低价高于上一日的收盘价，则在今日的最低价买入”时，就是一个很典型的含有未来信息的交易策略。在该策略中，买卖决策依赖于两个重要的价格信息：一个是上一日的收盘价；另一个是今日的最低价。上一日的收盘价显然是没有问题的，是一个从当日开始交易时就已经获得的信息。但是今日的最低价就是一个比较明显的未来信息或者说未来函数，因为不到今日收盘，是不可能确定当日的最低价的，要求在最低价买入是一个在现实中不可能完成的任务。不过在回溯测试过程中，由于已经在做决策时得到了关于当日最低价的信息，因此就有可能错误地将今日最低价当作可用信息并用来确定买卖决策。

还有其他一些关于未来信息错误使用的例子，比较知名的如“在牛市中持有股票”。实际上，除非具有预判牛市来临时间点的技术手段，否则这种说法往往只是牛市确认之后的事后诸葛亮。当一波牛市已经形成甚至走完的时候，再说出初期建仓之类的意见已经没有任何实际操作意义，而在牛市刚刚开始的时候，交易员常常无法准确地判定这到底是牛市的开始、持续的盘整还是前一波下跌途中的波动，也就无从谈起牛市持有股票了。

另外一个例子则更偏技术一些，当研究人员使用整个样本的数据进行优化并得到一个量化交易策略之后，使用该量化交易策略在全样本的整体回溯测试中判断买点、卖点、仓位大小等设置就属于引入了未来信息。这时的未来信息问题不在于量化交易策略本身，而在于量化交易策略的研究过程。用上面的示意图 3-1 来进行说明，就是优化量化交易策略时，无差别地使用了 A 阶段和 B 阶段的数据，但是最终优化得到的量化交易策略，又放到整个样本中进行回溯测试并得到了 A 和 B 之间的买卖点。因此在回溯测试的过程中，该买卖点的确认就使用了

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

B 阶段的信息，也就是未来的信息。

由于未来信息对于回溯测试的可靠性具有负面影响，因此在实际的策略研发过程中是需要极力规避的。最直接的办法莫过于将研发完成的量化交易策略放入实际环境中进行模拟交易或者实盘交易，原因在于实际交易中判断策略执行的时间点总是现在，因此必然是无法得到未来的信息的。如果在回溯测试过程中不慎加入了未来信息，具有相同结构的量化交易策略在实际的交易中也必将暴露出来。

但是在实际的策略研发过程中，由于量化交易策略会经历频繁的变更，因此全部采取实际检验的方法有时是难以达到的。在有些回溯测试的情形下，可以人为地将数据样本分成两个部分并分开存储，使用其中一个部分进行量化交易策略的优化工作，然后将优化完成的策略放入另外一个部分的样本数据进行模拟运行。这种强制分开储存数据的做法，其实也是在效仿实际运行中未来数据的绝对不可得。如果在实际的研发工作中数据分开存储也因为工作量过大而不易完成，那么似乎也只能依靠量化交易策略研究人员自身的逻辑思考和判别能力来规避未来信息了。

虽然这里给出了一些办法用来规避未来信息，但是仍然存在一些未来信息的种类是难以发觉甚至是难以避免的，例如在其他资料中讨论较多的存活者偏差。存活者偏差是指在某一个数据集当中，由于时间的推移，集合中之前包含的一部分成员由于无法存活而离开数据集，而在回溯测试时参照的是最后的存活成员，因此提前剔除了表现不好的非存活成员，造成了未来信息的问题。

依然使用图 3-1 来进行说明，研究人员在回溯测试时得到的数据是 B 阶段末尾所有存活成员的数据，但是在箭头所指向的买卖点时，所能见到的成员中还包括一些在 B 阶段离开的成员。回溯测试中由于数据集的问题直接把这些在 B 阶段表现不佳的成员进行了排除，就是滥用了 B 阶段才能得到的信息，从而虚假地提升了量化交易策略的回溯测试成绩。

这种问题常常发生在选股策略当中，一些股票数据库实际上已经根据当前的信息剔除了之前由于退市等问题而消失的股票的数据，使用这些数据库的数据进行回溯测试就会产生存活者偏差的问题。例如“买入价格极低的股票”这一策

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

略，在获得的数据具有存活者偏差的情况下，能够在测试中建仓的都是之后会存活下来的股票，在价格很难更低的前提下上涨的可能性很高，因此回溯测试的结果往往不切实际的优秀。一些指数、股票分类的调整等情况，也可能会导致存活者偏差的产生。

由于完全排除存活者偏差问题的数据库通常价格更高甚至无法得到，依靠研究人员在时间轴上自行收集数据组建数据库的做法又耗时耗力，因此在这类情况下完全规避存活者偏差是非常困难的。

3.2 过度拟合与欠拟合

在最简单的情况下，拟合是指对于一组已知的离散点，给定一个带未知系数的函数，通过调整该函数中系数的取值，使得已知的离散点与函数之间的差别最小化的过程。量化交易策略研发中的最优化部分可以在一定程度上看作一个拟合的过程，通过优化量化交易策略的模型设置和参数设置来让策略尽量适应交易资产的内在特征，从而获得更高的收益和更低的风险。

在这个最优化的过程中，可能会产生出一些偏离实际目标的结果，依照一般量化交易研究的习惯，作者称之为过度拟合⊖和欠拟合。实际上，在大的数据科学框架下，过度拟合和欠拟合具有更丰富的表现形式，下面的内容仅针对量化交易策略研究可能出现的情况进行说明，从而保证整本书内容的针对性。

这里用一个简单的例子来实际说明一下过度拟合和欠拟合。假设一个数据集源于一个带有误差项的二次方程，如下：

$$y=-x^2+15x+200+\varepsilon$$

其中 ε 是一个在 [−50，50] 之间均匀分布的误差项。数据集一共包括 20 个点，其中 x 分别选取从 1 ～ 20 的 20 个整数。如图 3-2 所示，星号为 20 个观测数据点，曲线为不带误差项的原始二次方程。

⊖ 往往被简称为“过拟合”。作者虽然也习惯于使用“过拟合”，但是为了行文准确，本书中统一使用“过度拟合”的说法。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

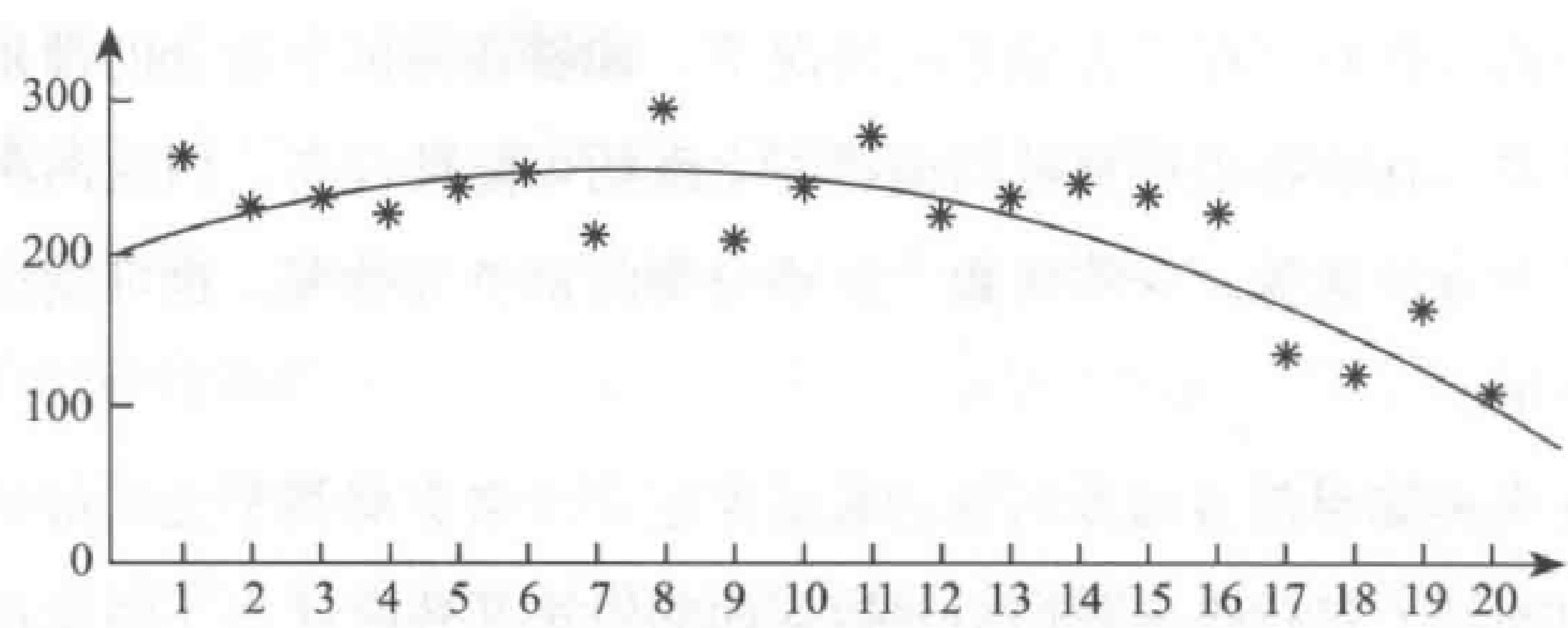


图 3-2 原始数据与来源

在实际工作中，研究人员并不知道数据来源的具体逻辑，因此可能会设置出各种各样的方法来拟合数据集。当研究者采用三次方程来拟合这 20 个数据点时，得到的方程曲线如图 3-3 所示。曲线与数据点的差距有一定缩小，但是曲线形状与图 3-2 中原始方程的曲线存在走势上的差别，在图形最左侧有一个原始图形没有的下降趋势，右侧的下降则比原图更甚。使用复杂度比二次方程更高的三次方程来研究该数据集，造成的就是过度拟合的问题。

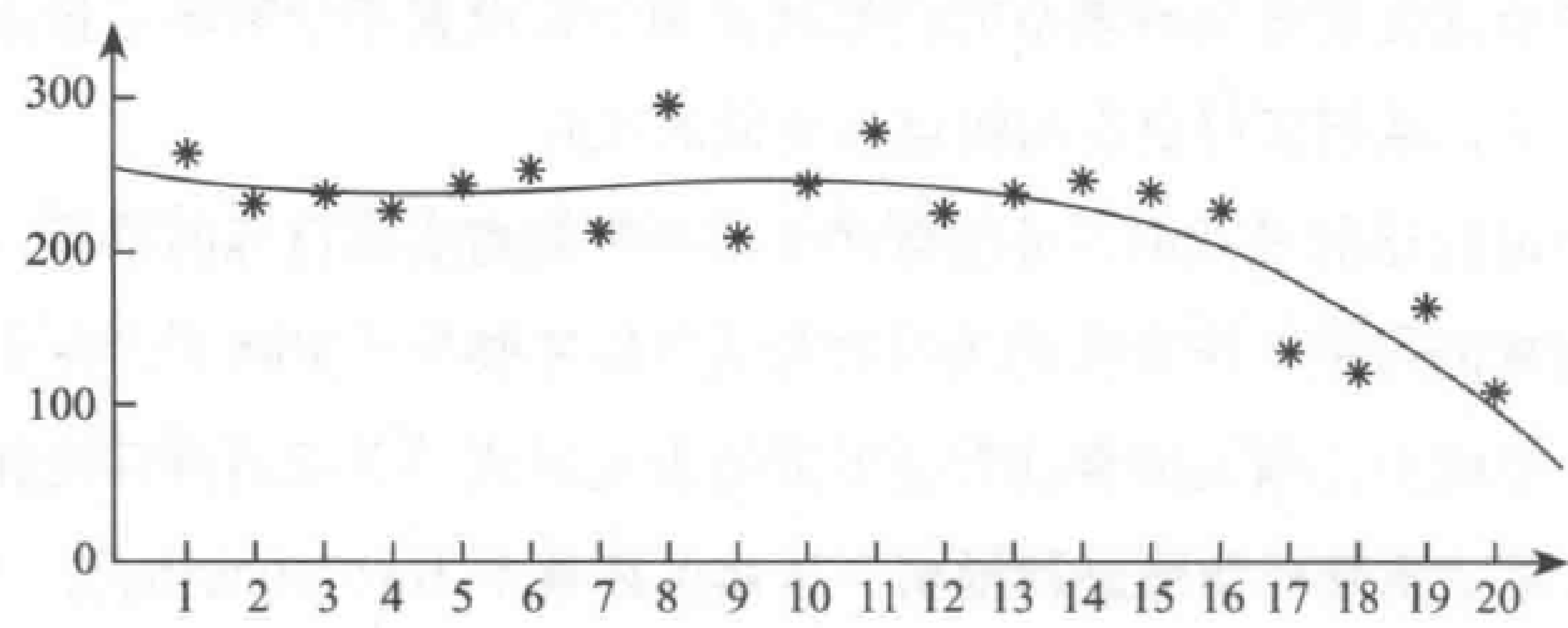


图 3-3 三次方程下的过度拟合

当研究人员使用一次方程来拟合这 20 个点的数据集时，得到的方程曲线如图 3-4 所示。很明显，用复杂度较低的线性方程拟合出的结果损失掉了原始曲线中先上升后下降的走势特性，而呈现出一个线性下降的趋势，这就是由欠拟合造成的。

上面这个例子简单化地展现了过度拟合和欠拟合在数据研究中的形式。对于量化交易策略而言，过度拟合指的是由于选取的策略或模型过于复杂，不仅适应了交易资产本身的内在属性，而且在优化过程中还刻画了噪声在样本数据中表

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

现出的一些特性，过度地拟合了样本数据，从而造成偏差，影响到策略的盈利能力。在一些专业书籍中，过度拟合问题有时候也被称为曲线拟合。欠拟合则恰好相反，往往是由于选取的策略或者模型过于简单，适应交易资产内在特征的能力不强，拟合水平较低，因此盈利能力也就较弱。

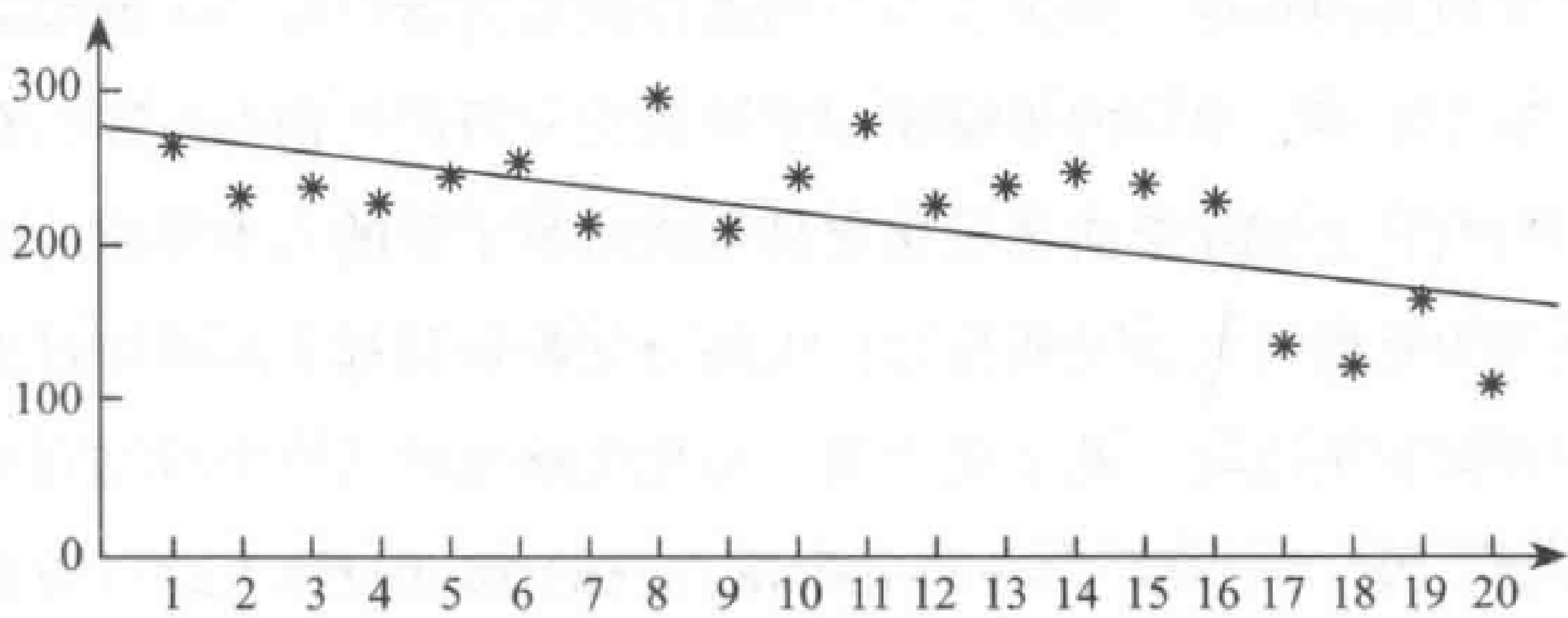


图 3-4　一次方程下的欠拟合

就过度拟合和欠拟合的具体表现而言，两者有相似之处也有不同之处。正如上面所述，两种问题都会导致策略的实际盈利能力下降，这是它们的相同之处。但是欠拟合的量化交易策略由于对数据的描述和挖掘能力不足，往往在历史数据优化下和实际交易中都表现不好，因此在回溯测试时就可以比较直观地分辨出来。过度拟合的量化交易策略则不同，由于在拟合样本数据时过度优化，进而拟合了样本中噪声的特性，因此常常是历史数据优化下的结果很好，只有在实际交易中才会显著变差，也就是俗称的泛化能力低下。在实际的量化交易策略研发过程当中，这样的性质增加了过度拟合判别的难度。

在大部分情况下，一个量化交易策略往往既具有过度拟合的问题，也具有欠拟合的问题，原因在于金融资产的内在属性并不像物理学等学科所研究的对象一样，存在一个较为明确的运行逻辑，同时也比语音识别等数据科学问题的情况更加复杂和模糊，人们现在对于资产价格运行模式等问题的了解还处在非常浅层的状态。在实际工作中，研究选取的策略或模型可能只有一部分设置适应了交易资产的内在属性，相符的内在属性也只占交易资产整体属性的一部分，因此策略模型是欠拟合的。好在量化交易策略研发并不是一个需要绝对精确的学科，即使策略模型与数据内在特性相符的程度不大，只要存在这种适应性，就有可能为研发

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

者带来可观的利润。与此同时，研究得到的策略模型的另外一部分设定可能就与交易资产的内在属性完全没有相关性，经过最优化等工作之后只是拟合了资产交易中的噪声，从而造成过度拟合的问题。

实际工作中，研究人员既需要减少量化交易策略过度拟合的程度，也要同时减少策略欠拟合的程度。解决欠拟合的途径较为清晰，就是不断地尝试各种各样的策略，基于更多、更好的策略模型来挖掘交易资产的内在特征，获得相应的盈利能力。当然，尝试新策略是量化交易策略研究工作的基本要求，因此欠拟合这一问题也较少地被研究人员所提及，大部分文献资料都更注重于过度拟合的问题，也存在许多相关论述。从本质上说，简化策略当然是一个行之有效的较少过度拟合的方法，冯·诺伊曼（John von Neumann）曾通过他人之口说出了“我能用四个参数拟合一头大象，用五个参数让它摆动鼻子”的经典名言，参数过多、模型过于复杂可能会导致过度拟合这一看法已经成为一个公论。

但与此同时，简化策略又与研究者要解决欠拟合、尝试找到更合适策略的目标背道而驰。图 3-5 给出了一个策略在简单和复杂之间变动时，欠拟合与过度拟合情况的粗略示例。由于交易资产的内在运行逻辑尚不清楚，因此任何一个交易策略都只能挖掘并利用数据内在特征的一部分，也就是两个圆重合的浅灰色区域，而剩下的白色区域则是交易策略没有实际效用但是客观存在的部分。当策略复杂度增加时，策略就有可能更多的利用数据的特征，表现为浅灰色的重合区域增加，欠拟合问题得到缓解。但是同时，策略无效的白色区域也可能相应地增加，这一部分经过最优化就是过度拟合的成因。



图 3-5　策略复杂化的示例

因此，在量化交易策略的研发工作中，策略的复杂程度、参数数量的设置等问题常常需要研究人员根据实际情况、过往经验来进行主观判断，没有一个固定

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

的最优标准。好在针对回溯测试过程而言，存在着一些可以用来判别过度拟合问题的技巧，因此研究者可以在适当增加量化交易策略复杂程度的基础上，通过这些技巧来尽量规避策略的过度拟合，本书将在第 6 章中介绍的推进分析正是一种处理过度拟合问题时行之有效的分析技巧。

除此之外，增加回溯测试时使用的样本量也是一个可能会产生作用的做法。如果量化交易策略所针对的交易资产内在特性没有随着时间产生变化，那么增加样本量可以扩展回溯测试的覆盖范围，从而在更大的区间内研究策略的稳定程度，而样本量的增加也可以让研究人员更自由地使用一些判断过度拟合的研究技巧，这样就能够在保证不造成欠拟合问题的条件下，更好地规避过度拟合的产生。在实践中增加样本量进行研究的做法也确实能够取得一定的效果，因此作者建议在可获得的数据量充足的前提下，还是应该尽可能地让回溯测试覆盖更广的范围。

不过究其根源，内在特性不随时间变化的假设本身是值得怀疑的，前面已经说过，人们现在对于资产价格运行等内在特性的了解还处于非常浅层的状态，即使研究者发现了某些规律并基于这些规律形成了可以盈利的策略，也不能保证这些规律会保持下去。在这种情况下，不经判断而盲目地增加数据量，就可能对量化交易策略的盈利能力造成误判。而且在中国这种运作时间较短、环境变化较快的市场上，该问题就更加严重。完全消除这个矛盾需要研究人员对本质问题充分理解和把握，这并不是仅靠一个合理的量化交易策略研发框架就可以解决的问题。

3.3 回溯测试与真实环境的差异

回溯测试中的交易环境与真实的交易环境是有一定差异的，这种差异越大，回溯测试的结果用于判断策略效用的能力就越低下。因此，理解回溯测试与真实环境之间的差异问题，学习如何处理两者之间存在的差异，也是本书在讨论具体的量化交易策略之前需要讲解的内容。

回溯测试与真实环境之间差异的情况，大致可以分为两类。其中一类差异是

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

可以通过人为设置回溯测试或者人为限定真实交易的手段而基本消除的。在这一类差异当中，实际交易时的买卖点与回溯测试下的买卖点不一致的问题是一个比较常见的例子。

具体来说，就是在中低频的量化交易策略研发中，回溯测试使用的数据往往不是价格进行一次变化就记录一次的分笔数据，而是记录一段时间内的开始价格、最高价格、最低价格、结束价格之类的二次处理后的数据。因此，如果真实交易中的买卖点发生在开始价格和结束价格之间的某一个时间点时，是不能反映在二次处理数据所构造的回溯测试中的，也就会造成这两者买卖不一致的问题。

图 3-6 给出了一个交易日内买卖与回溯测试不符的实际案例。假设在这个例子中，研究人员在回溯测试中使用的数据是日度收盘价所组成的时间序列，测试的量化交易策略为“当价格超过 10 日均价时买入，当价格低于 10 日均价时卖出”。图 3-6 中给出的是某一交易日全天的价格走势，假设这一天的交易价格达到 53.5 以上时，则触发“价格超过 10 日均价”的条件需要买入。

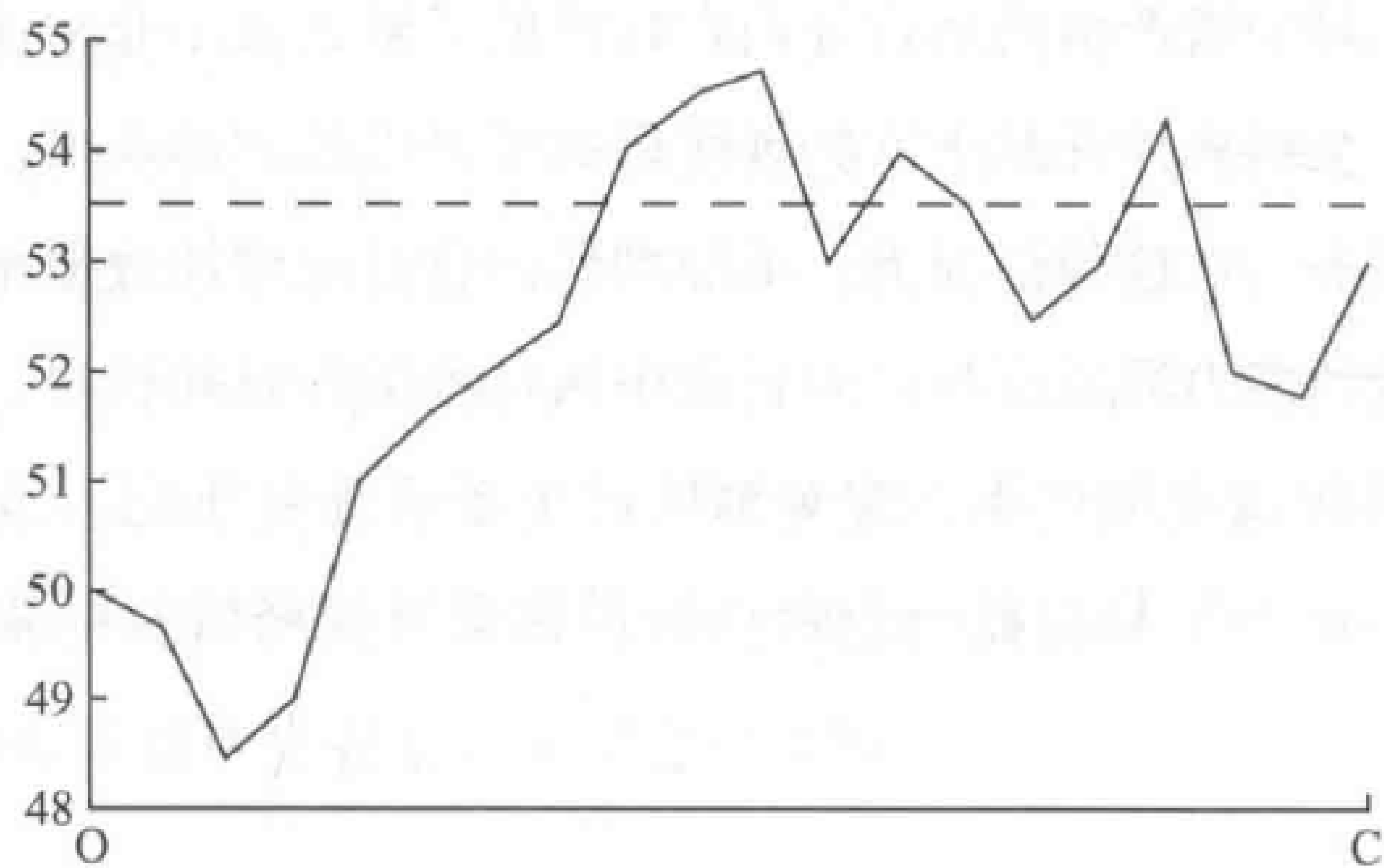


图 3-6 一个交易日内买卖与回溯测试不符的实际案例

从图 3-6 中可以看到，在回溯测试的流程中，由于当日价格报收于 53，没有达到 53.5，因此回溯测试时是判断该日没有交易行为的。但是查看实际的全天价格走势，可以发现价格一共三次突破了 53.5，并又三次下跌到 53.5 以下，如果采用回溯测试得到量化交易策略并全天候地应用于资产交易中，这一天其实已经买入并卖出了该资产三次了。

如果这里的资产是 A 股中交易的股票，那么情况可能更糟，在交易日内第

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

一次超过 53.5 时进行了买入操作之后，由于 T+1 的规定不能再在该日卖出，最后的结果是回溯测试中该日没有交易，但是实际上该日收盘后持有了股票仓位，这种差异将完全破坏掉回溯测试结果的可信程度。

对于如何处理这一类的差异，作者给出以下三种可供参考的解决方案。

（1）在回溯测试时，使得回溯测试的环境尽量与真实的交易环境保持一致。在上面的例子中，就是将回溯测试使用的二手数据换成更为精确的分笔数据，从而模拟整个交易过程中都可以买卖的现实环境。

（2）在实际的交易过程中，尽量贴近回溯测试中的交易情况。以上面的例子而言，就是在实际交易中忽略每一日盘中的走势，只在收盘时判断交易与否。但是值得注意的是，由于得到收盘价的同时也就代表着当日交易的结束，完全按照收盘价进行交易决策会导致不能执行的问题，因此这种解决方案需要在实际收盘点提前一小段时间进行交易决策并执行。

（3）合理估计实际环境下的交易与回溯测试之间的差异，并在回溯测试过程中通过计算上的调整反映出这些差异。在上面的例子里，实际上就是估计出相对于回溯测试中的收盘交易而言，盘中交易会带来的收益和损失，并在回溯测试里加入对估计收益和损失的刻画。但是如果实际差异涉及 A 股市场 T+1 造成的建仓不符之类的问题，那么这种解决方案可能就不是一个很好的选择了。

此外，改变量化交易策略，使之能够适应这种回溯测试与真实环境之间的差异，也是一种在实际策略研发过程中可能会用到的处理手段。比如上面的例子中，可以在原始的量化交易策略判断买卖点时，加入合适的限定条件或者缓冲条件，从而解决频繁地来回买卖的问题。但是这种方法治标不治本，即使来回买卖的问题得到了缓解，但是回溯测试与真实环境之间的差异仍然会存在，如果希望将这种差异消除掉，可能仍然要结合上面所述的三种方案或者其他方案来完成。

对于这三种方案而言，作者建议在有能力的情况下，优先考虑第一种解决方案，即使很难获得相同样本期下的分笔数据，也尽量采用频率更高的数据来代替之前的数据，从而更好地逼近真实的交易环境。随后的选择才是第二种和第三种解决方案，这两者之间没有一个固定的更优选择，需要结合具体的量化交易策略、研究环境和交易环境进行人为判断。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

当然，也存在着其他一些具有实际意义的操作方案，例如将第一种和第二种解决方案结合起来，同时更新回溯测试和实际环境中的交易行为。再以图 3-5 中所展示的例子来说明，完全在收盘时进行判断和买卖在现实环境中是不可能的。而由于在收盘之前的短暂时间进行交易的判断和执行稳定性较差，而且用来判断交易的数据点也不是收盘价，因此会存在较大的差异。在频率更高的样本数据无法获取的情况下，其实可以在回溯测试中改用开盘价进行研究，如果测试结果也可以得到合适的收益和风险情况，那么执行时就可以在开盘之后立即执行，这种情况下所使用的判别点就是当日的开盘价，差异仅存在于开盘和执行之间的时间差，而且这种差异较之前一种情况来说，一般会更小一些。

需要说明的是，在本书后面介绍的所有量化择时策略与量化选股策略案例中，使用的都是收盘价等收盘时获取的数据。也就是说，回溯测试的模拟交易过程，都默认将收盘时刻作为进行交易决策的时间点。因此，对实际环境下的交易与回溯测试之间的差异进行估计从而判断出策略的实际效用，就成为一个本书没有具体分析但是实际上非常重要的研究组成部分。

在现实研究中，将本书所介绍的几个简单的交易策略全部更改为开盘时刻进行交易决策并不困难，相关的数据也比较容易得到，对此感兴趣的读者可以自行尝试。但是由于本书的目的是介绍量化交易策略研发的流程框架，实际案例仅作为辅助介绍的工具使用，因此也就依照一般的研究传统，将数据确定为收盘数据。

此外还可以额外地告诉读者，将自回归策略改为开盘时刻获取数据、计算模型并进行交易之后，本来盈利水平很高的策略不再具有显著的盈利能力。推进分析甚至多层推进分析没有检验出来的问题，在改变样本数据集之后得以呈现。因此这里再次申明，作者希望读者在阅读本书时，不要将关注的重点放在具体策略之上，本末倒置，熟悉研究框架、培养研究思路才是最为重要的。

另外一类回溯测试与真实环境之间的差异则是很难通过人为的修改和设置来消除的，其中最常见的是实际环境中会产生的交易成本。估计交易成本是量化交易策略从研发进入实际执行前一个必要的步骤，本书在第 2 章的研发流程中曾经谈到了交易成本问题，在后面的第 14 章中则会用一章的内容对其进行分析，因

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

此这里不做过多阐述。

流动性也是一个时常会导致回溯测试产生重大缺陷的问题，这个问题在 A 股市场上同样格外明显，到目前为止，国内的 A 股市场仍然设置涨跌停板，如果某交易策略判断买入一只将会涨停的股票，在回溯测试中可以正常建仓，但是真实交易中可能由于当日的开盘时刻就涨停，因此实际上是没有机会买入这只股票的。将回溯测试使用的二次处理数据更换为更精确的数据，如分笔数据或者分钟数据、小时数据等来进行研究，可以在一定程度上减小这种情况下回溯测试与真实环境之间的差异。

当然这只是流动性缺失的一个比较极端的例子，在大部分存在流动性问题的情况下，可能会出现的结果是实际交易价格与回溯测试的虚拟交易价格不一致，研究者通常将这种差异也归为交易成本的一种，并通过预估来尝试解决。总的来说，处理这一大类的差异主要还是采用前面提到的第三种方法，即合理估计实际环境下的交易与回溯测试之间的差异，并在回溯测试过程中通过计算上的调整反映出来。

此外，很难通过人为的修改和设置来消除的差异还包括量化交易从业者对于量化交易系统的整体判断和情感因素。诚然，量化交易策略的优点之一就是可以规避策略研发者的主观臆断、排除情绪化对于整个交易过程的干扰和影响，但是研究和控制量化交易策略的毕竟还是人，因此总体上的行为偏误依然会存在。

例如，量化交易员对于量化交易策略整体风险的承受能力，虽然在研发过程中，研究者可能会设定一个 30% 的警戒线，如果亏损达到 30%，则判定策略失效并放弃策略，但是在实际的运行过程中，5% 的亏损已经让人惴惴不安，10% 的亏损则让人开始怀疑策略的正确性并重新检验，20% 的亏损可能就让人放弃坚持并修改策略了。有主观判断存在，就有偏误和差异存在，这是交易策略都固有的问题。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

第 4 章

简单的择时策略

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

4.1 择时策略的基本框架

在这一章中，作者将向读者介绍最基本的择时策略是如何构成的，以及相应的一些实际案例。首先需要说明，这里所谓的最基本的择时策略，指的是只判断买卖、不涉及仓位优化的择时策略，图 4-1 通过灰色区域说明了基本的择时策略在研发流程中所处的位置。由于不考虑风险的问题，因此择时策略完全是通过优化收益情况来形成相应的买卖决策的。

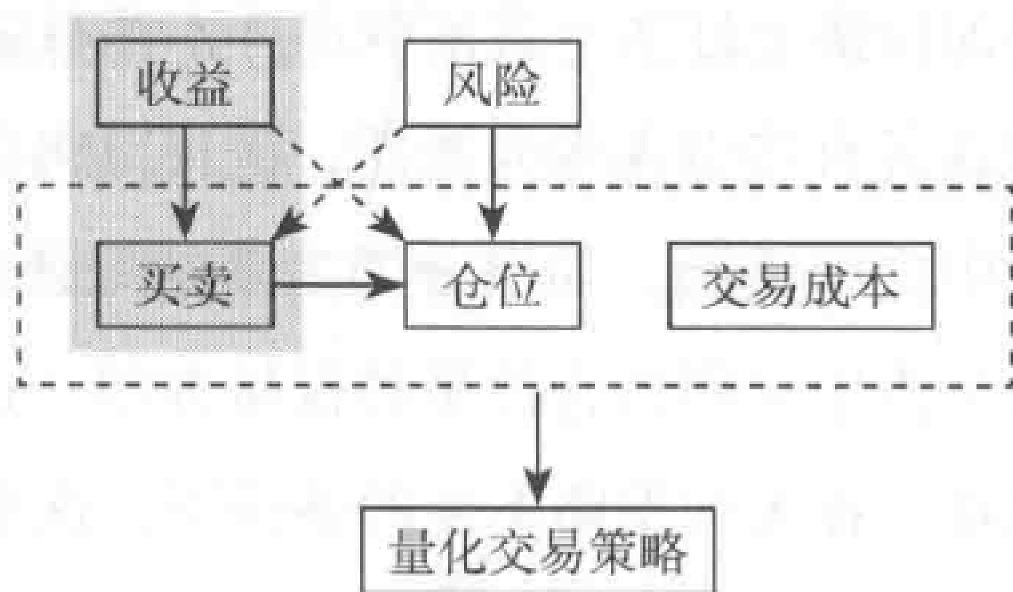


图 4-1 基本的择时策略在研发流程中的位置

另外需要注意的是，这一章的主要目的是为读者提供较为清晰易懂的量化择时策略案例以供入门，因此会在策略的优化过程中使用全样本数据进行设置的优化，并直接使用最优化下的买卖决策作为模拟交易的结果加以分析，这也是章名称取为“简单”的意思。这种做法会产生出未来信息的问题，希望读者暂时忽略，后面的章节中会有相关内容来处理这一问题。

一个基本的量化择时策略框架如图 4-2 所示，无论是回溯测试还是实际交易，每一次做出交易决策时，都根据决策时刻的已知数据，通过策略构建的分类器或分类函数将当前的状态分为两类。第一类，在分类为多的情况下，如果当前仓位为多头，则保持不变；如果不是，则通过买入操作将仓位变为多头。第二类，在分类为空的情况下，如果仓位为空头，则保持不变；如果不是，则通过卖出和卖空操作将仓位变为空头。另外，假如市场禁止做空，那么在分类为空的情况下，如果仓位为空仓，则保持不变；如果不是则通过卖出操作将仓位变为空仓。

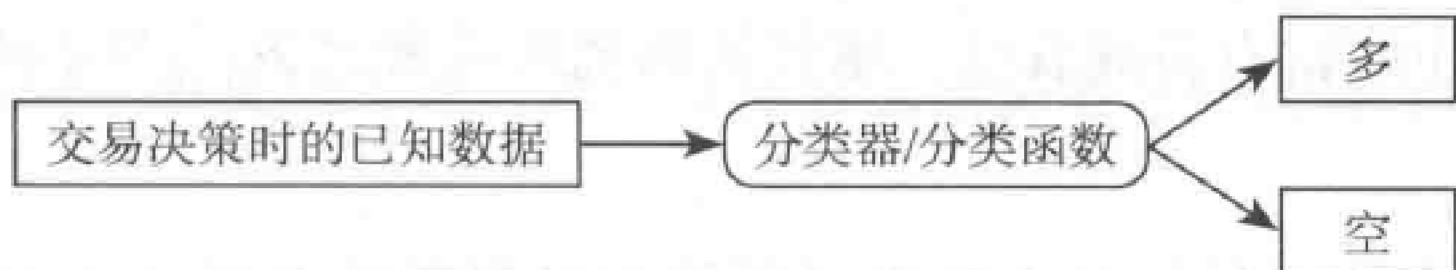


图 4-2 基本的择时策略框架

这是一个相对而言比较简单的过程，期间不涉及对交易决策点之后收益情况的预测，交易决策仅仅和当前状态的分类有关，因此很多基于这种模式构造量化

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

交易系统的从业者声称自己只跟随系统、不做预测。不过在作者看来，分类的做法本身暗含着对未来状态的分类，这种情况类似于一个条件预期的计算和判断，预测隐藏其中。当然，怎么理解一个框架是见仁见智的问题，并无优劣之分，量化交易策略的研究人员只需要明确自己的理念和判断即可。

有时候一个量化交易策略的使用者在整个交易的过程中并非只有买和卖这两种选择。例如在市场允许卖空的情况下，既不买也不卖、持币观望也是一种很重要的做法。好在分类函数也并不是只能处理二分类的问题，通过组合多个二分类器或者直接引入多分类器，就可以将分类结果扩展为多种。图 4-3 的框架就加入了中立的类别，如果分类为多就保证仓位为多头，分类为空就保证仓位为空头，分类为中立则将仓位平掉保持观望。这实际上是图 4-2 中框架的一个顺理成章的演化，在大的思路不变的条件下，这个框架还能够进一步的复杂化，从而适应具体的交易问题。

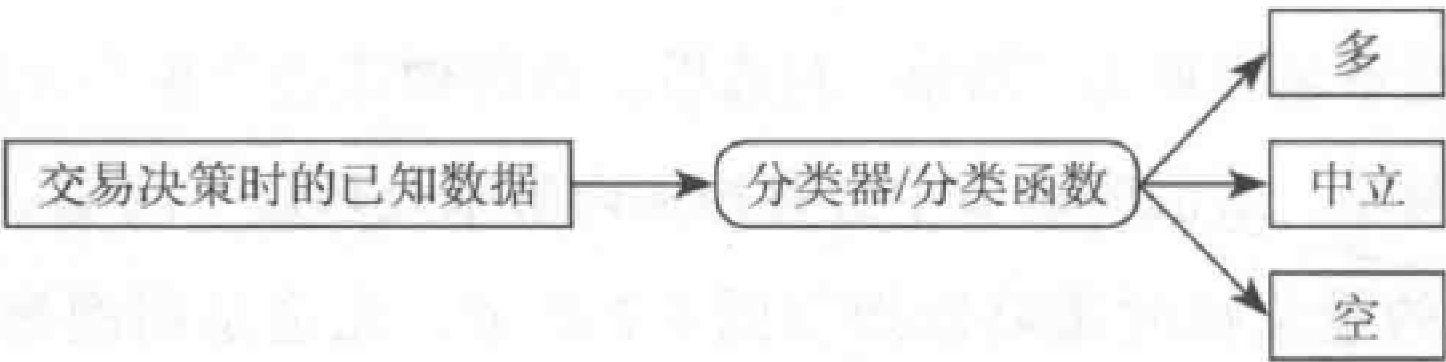


图 4-3　多分类的基本择时策略框架

在图 4-2 和图 4-3 的量化交易策略框架中，分类器 / 分类函数的含义是比较广的。例如一个技术指标的使用，当该技术指标符合某一种判别形式时则看多，符合另一种判别形式时则看空，就是一种很典型的分类。对外部环境的判断也可以归于这样的框架之下，当交易资产的外部环境向好时则看多，当外部环境变差时则看空。对新闻事件的影响进行押注的交易策略也可以看作一种分类，当某个特质的事件发生时就看多或看空，事件影响消失或者过去一段时间之后则分类回中立状态。

当然，只有使用量化的方法来对环境加以判断或者对事件加以处理，才能将其置于量化交易策略的框架之中。实际上，大部分的择时策略都可以纳入这样一个框架中，更复杂一些的择时策略组成成分，如过滤器之类的辅助判别工具，可以看作分类技术的进一步细化。而仓位建立之后的进一步加仓、减仓、止

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

损、止盈等手法，则可以看作分类类别更多的分类器所导出的结果。当然，在现实的量化交易策略研究当中，有时候把加仓、减仓当作另外一次开仓、平仓的过程来对待，然后将多次交易行为合并起来，可能会是一种更为高效的处理方法，逻辑上也更为清晰一些，不过这并不妨碍我们使用上面所述的框架来理解整个策略。

基本的量化择时策略还有一种比较特殊的形式，即基于对交易资产未来收益的预测来形成具体的交易策略，一个简单的框架如图 4-4 所示。图 4-2 和图 4-3 中的基本择时策略在表面上均不涉及对交易决策点之后收益情况的预测，交易决策仅仅和当前状态的分类有关。图 4-4 中的择时策略框架则不同，主要的工作往往就是在预期收益的计算和处理上。

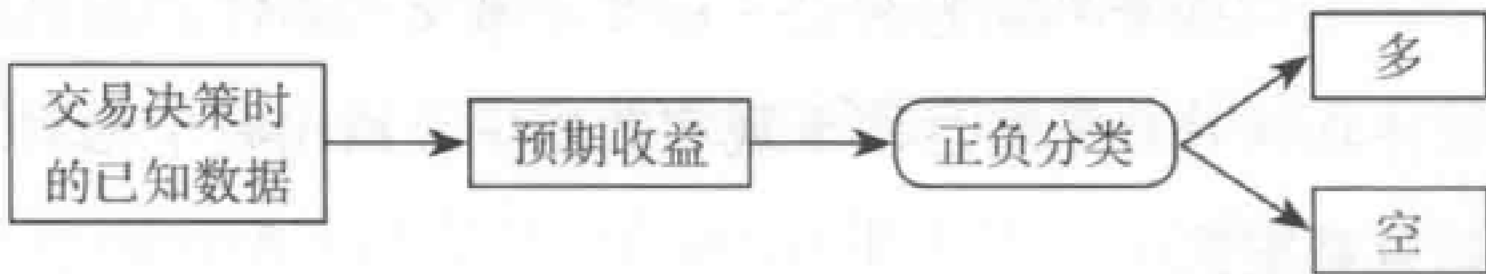


图 4-4　基于收益预期的基本择时策略框架

由于这里讲解的基本择时策略不涉及风险的估计和仓位的设置，因此最后要做出的交易决策只包括买和卖。图 4-4 中就采用了一个简单的正负分类来判断多或空，如果预期收益为正则看多，如果预期收益为负则看空，买卖决策方面与图 4-2 中的情况相似。当然，也可以通过设定零上下的阈值来形成多种分类，如同图 4-3 一样判断出多、空、中立等类别，从而做出相应的交易决策。

需要注意的是，对于这一类的策略框架而言，由于存在着预测未来收益这样一个中间步骤，因此在参数优化时既可以选择以收益预测模型拟合的最好为优化目标，也可以选择以整个策略最后得到的交易结果最优为优化目标。在实际的研究中一般都采用前一种优化方式，当然具体的选择仍需要针对具体的研究问题来加以判断。

从上述几种框架的特性出发，其实就很容易理解目前机器学习的一些算法在量化交易策略中不断应用的趋势了。机器学习和人工智能领域中一些在应用上较为成熟的算法，很多都是以分类为目标的，同时一些算法也可以用来进行非线性回归之类的拟合和预测，因此本质上就和量化交易策略的框架具有相容性。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

但是即便如此，作者仍然不太认同行业中存在的拿来主义，一些量化交易策略的研究人员在没有搞清楚机器学习算法内在逻辑的情况下，简单地将现成的工具包直接套用到交易数据之上，其实是不可取的。例如使用支持向量机做分类，至少也应该先弄清楚使用支持向量机来判断类别所具有的优点和缺点，是否适用于自己的策略和实际情况。况且一部分机器学习算法在工业界的应用是有大规模的样本数据作为训练基础的，直接用来构造量化交易策略时，会有多大的可能性造成过度拟合，都是应该提前考虑的问题。

4.2 均线趋势策略的简单优化

使用资产价格的均线来判断趋势是一种非常常见的做法，而且量化起来简单直接，这一节就将均线趋势策略运用于现实数据中，给出一个相对来说比较容易上手的量化择时策略案例。

图 4-5 在图 4-2 所示的框架基础上，说明了一个双均线趋势策略是如何运行的。在交易决策的时点需要根据已知数据计算短期均线和长期均线两个值，当短期均线高于长期均线时，判断交易决策时刻的趋势为上涨趋势，按照趋势会延续的思想，认为后市会继续上涨，因此看多。而当短期均线低于长期均线时，判断交易决策时刻的趋势为下跌趋势，认为后市继续下跌，因此看空。

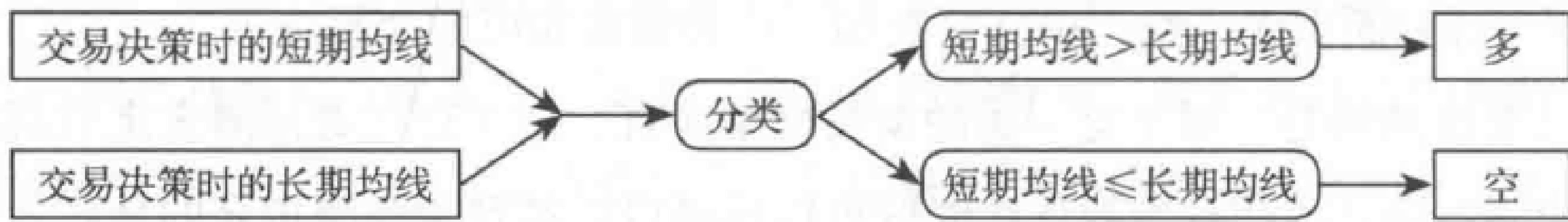


图 4-5 均线趋势策略的基本框架

图 4-5 中的两个分类标准一个是大于，一个是小于等于，主要是因为最后分出的类别仅为两类，因此要保证两个分类标准合并起来涵盖所有的可能性，等于号出现在上一个或下一个标准当中对实际交易情况会有一定影响，但不会太大。

前面已经提到过，这一章的目的主要是为读者提供较为清晰易懂的量化择时策略以供入门，因此会在策略的优化过程中使用全样本数据进行一次性的优化，从而使整个研究过程简单化。以这一节的均线趋势策略为例，也就是在固定了均

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

线趋势策略的基本框架之后，对于短期均线长度和长期均线长度两个参数，在全部模拟交易样本上考虑不同参数组合所带来的收益情况，选取收益最好的一个参数组合为优化后的参数组合。随后考察这一个参数组合的均线趋势策略在相同样本上的模拟交易表现，进而判断均线趋势策略的优劣。

实际上这里通过整体优化引入了未来信息，作者将在后面的章节中解决这一问题，本节则直接考察相关结果。由于策略构建过程中，在全部样本数据上只优化了一次，也就是说模拟交易样本下都将使用同一对参数组合，具有较好的一致性，因此优化确定了参数组合的均线趋势策略实际上等同于“短期均线从下向上穿过长期均线则做多，短期均线从上向下穿过长期均线则做空”。

这个案例中选取的交易资产为沪深 300 股指期货主力合约，其标的指数为中证指数公司编制的沪深 300 指数，由上海和深圳证券市场中规模大、流动性好的 300 只股票组成。沪深 300 指数期货是中国国内最早上市也是最有影响力的股票指数期货，由中国金融期货交易所在 2010 年 4 月 16 日推出。本着尽可能多地利用数据的原则，这里获取了从 2010 年 4 月 19 日到 2015 年 6 月 30 日的日收盘价数据，判断买卖则从 2010 年 12 月 31 日收盘开始，也就是说在回溯测试样本中将一共模拟交易四年半的时间。2010 年 4 月 19 日到 2010 年 12 月 30 日的日收盘价数据作为交易前计算短期均线和长期均线的备用数据，数据长度为 172 个交易日，也就是说只要长期均线在 173 日均线以下，备用数据都可以支持均线趋势策略的计算顺利完成。虽然股指期货可以在最低交易保证金为 8% 的情况下进行杠杆交易，但是在本节的基本择时策略环境下，统一将杠杆率设定为 100% 进行研究。

这里进行的优化，其实就是很简单地将所有参数的可能组合都测试一遍，然后选取整体收益最高的参数组合，这种方法也被称为网格搜索。但是对于一个均线策略而言，考察从 1 开始直到某一个长度结束所有可能的均线情况其实是一种存在冗余和偏向的做法。长度越大的两根相邻均线之间的差异越小，例如 2 日均线和 3 日均线之间的差异，远远大于 100 日均线和 101 日均线之间的差异，后两者由于一个交易日的差距被百分之一化，因此相差其实很小。这样，如果从 1 开始直到某一个长度结束所有可能的均线均参与搜索，那么长度小的均线的变化较

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

快，搜索会偏向于变化不大的长期均线，进而造成计算上的冗余。

前人采用过的改进方法包括将所有的短期均线加入搜索范围，长期均线则每5个步长加入搜索等，本书则使用斐波那契数列来处理均线的搜索范围。对于这个例子而言，斐波那契数列并不是一个具有神奇效果的技术工具，仅仅是作者觉得该数列在数学上可以比较好地将均线之间的差异平均化而已，换成其他的抽取形式其实也是可以的，请读者不要过分联想。斐波那契数列的形式为 [1，2，3，5，8，13，21，34，55，89，144，…]，也就是说，均线趋势策略的简单优化将使用1日均线[⊖]、2日均线、3日均线、5日均线……144日均线一共11种均线的长短组合进行搜索，寻找盈利能力最好的组合，从而确定最优的均线趋势策略形式。这里限定短期均线最大搜索到34日均线，长期均线则从短期均线的后一根开始搜索一直到144日均线结束（见表4-1）。

表 4-1　所有组合下均线趋势策略的年化收益率　（%）

长期均线＼短期均线	1	2	3	5	8	13	21	34
2	0.33							
3	0.22	8.60						
5	−0.21	8.08	−15.17					
8	5.48	−3.31	−6.25	2.72				
13	5.65	10.32	8.06	10.02	7.68			
21	19.29	18.90	23.53	16.77	18.95	18.11		
34	16.00	17.08	18.50	18.43	18.58	6.54	−2.92	
55	7.59	13.84	14.60	14.35	7.37	7.75	−0.93	0.28
89	12.42	13.85	11.10	15.33	11.82	13.04	13.24	11.67
144	8.22	13.83	15.35	11.00	8.69	3.46	−4.06	1.83

表4-1给出了所有的短期、长期均线组合下均线趋势策略的年化收益率情况。整个表中年化收益率为正的均线趋势策略一共有45个，年化收益率为负的仅有7个，因此纯粹从优化的角度来看，均线趋势策略是一个相对而言比较稳健的策略。收益率表现最差的一个策略是短期均线为3日均线、长期均线为5日均线的趋势组合，年化收益率为 −15.17%。收益率表现最好的一个策略是短期均线为3日均线、长期均线为21日均线的趋势组合，年化收益率达到23.53%。值

⊖ 实际上就是日收盘价序列。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

得注意的是，这两个均线组合中的短期均线都是 3 日均线，也就是说在使用 3 日均线形成的短期均线时，均线趋势策略的收益表现具有比较大的变化幅度，这在网格搜索当中是一个不太好的现象。但是同时，长期均线是 21 日均线的 6 个均线趋势策略的收益都比较高而且相对稳定，年化收益率最低的一个也达到了 16.77%。综合来看，通过收益最优原则选择出来的短期均线为 3 日均线、长期均线为 21 日均线的均线趋势策略，所处的区域应该算是比较稳健的。

在优化完成之后，再将短期均线为 3 日均线、长期均线为 21 日均线的均线趋势策略应用到选取的模拟交易样本之上进行回溯测试，从而得到回溯条件下相应的交易情况。图 4-6 展示了使用这一均线组合的均线趋势策略进行交易的情况下，以 1 为初始值的净值走势情况。在 2011 年年底、2012 年年底、2013 年年底和 2014 年年底，交易策略下的净值分别变为 1.3234、1.4452、1.4163、2.1586。由此可以发现，该均线趋势策略在 2011 年的收益情况较好，随后进入了一个为期将近三年的平缓上升。从 2014 年 11 月开始直到 2015 年 6 月样本期结束一共 7 个月的时间里，净值水平有了一个明显的提升，从 1.5 左右上涨到结束交易时的 2.5878，涨幅大约为 72%，当然期间的波动也较为明显。基于整个交易过程来看，四年半的策略收益率为 158.78%，折算下来的年化收益率达到 23.53%。

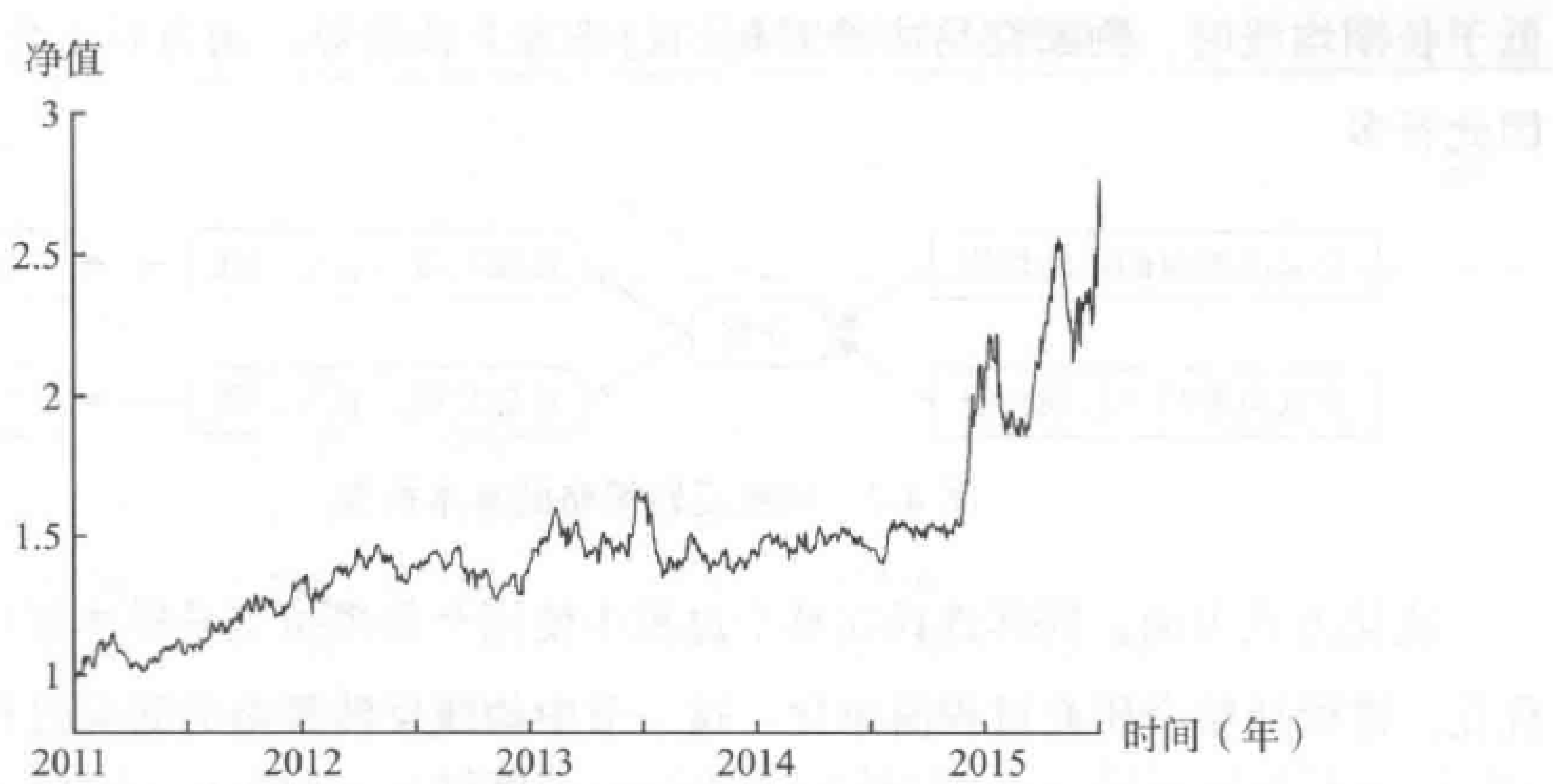


图 4-6 简单优化的均线趋势策略净值走势

在整个四年半的均线趋势策略交易中，净值的最大回撤率为 18.59%，而且从 2011 年 1 月 10 日收盘开始，就再也没有下降到 1 以下的水平。最大回撤的

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

发生时间在2013年中，最高点为2013年6月25日收盘的1.6528，最低点则为2013年8月1日收盘的1.3456，整个回撤过程一共经历27个交易日，是一个相对而言不算太长的回撤时间。经过计算，整个交易过程中平均的年度最大回撤率为13.87%，对比23.53%的年化收益率，该回撤水平是可以接受的。

4.3 均线反转策略的简单优化

均线反转策略与均线趋势策略是一对具有相反内涵的量化择时策略，同均线趋势策略一样，均线反转策略同样具有量化起来简单直接的特点。在这一节中，作者将均线反转策略基于同样的研究方法，运用到相同的数据样本之中，一是用来进一步介绍简单的择时策略，同时也能够与上一节的内容进行参照和对比，促使读者形成灵活运用量化工具的思维。图4-7在图4-2的框架基础上，说明了一个双均线反转策略是如何运行的，与图4-5对比起来看，其实就是将分类步骤中的大于号和小于号进行了调换。在交易决策的时点需要根据已知数据计算短期均线和长期均线两个值，当短期均线高于长期均线时，判断交易决策时刻的趋势为上涨趋势，按照走势会反转的思想，认为后市会下跌，因此看空。而当短期均线低于长期均线时，判断交易决策时刻的趋势为下跌趋势，认为后市会转为上涨，因此看多。

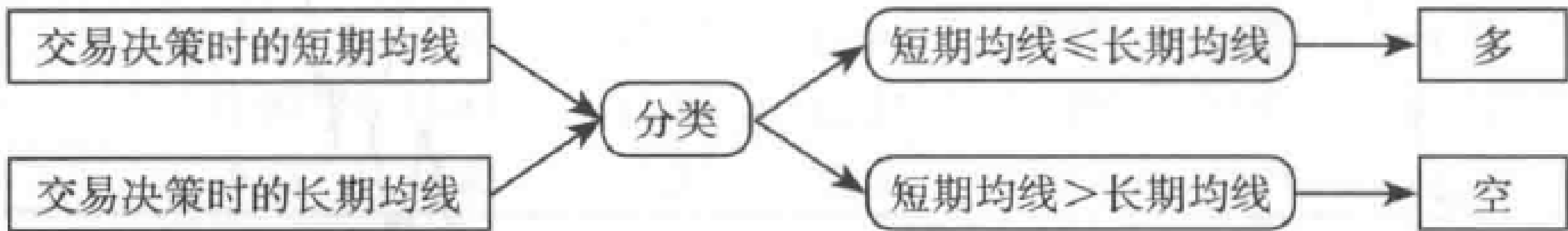


图4-7 均线反转策略的基本框架

优化方式方面，同样选择在整个过程中使用全部模拟交易样本进行一次性的优化，进而让整个研究过程简单化。这一节中均线反转策略的研究过程基本上就是上一节均线趋势策略的翻版，因此不多做赘述。由于策略构建过程中，在全部样本数据上只优化了一次，也就是说模拟交易样本下都将使用同一对参数组合，因此优化确定了参数组合的均线反转策略实际上等同于“短期均线从上向下穿过长期均线则做多，短期均线从下向上穿过长期均线则做空”。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

选用的交易资产依然为沪深300股指期货主力合约，具体的样本数据是从2010年4月19日到2015年6月30日的日收盘价。判断买卖从2010年12月31日收盘开始，也就是说在回溯测试样本中均线反转策略一共模拟交易四年半的时间。2010年4月19日到2010年12月30日的日收盘价数据作为交易前期计算短期均线和长期均线的备用数据，数据长度为172个交易日，也就是说只要长期均线在173日均线以下，备用数据都可以支持均线反转策略的计算顺利完成。股指期货交易的杠杆率也同样被设置为100%，从而使策略的研究过程相对简单。

在进行网格搜索的优化过程中，也依然使用斐波那契数列来处理均线的搜索范围，用以缓解搜索全部均线时可能会出现的计算冗余和搜索偏向的问题。斐波那契数列的形式为[1，2，3，5，8，13，21，34，55，89，144，…]，也就是说，均线趋势策略的简单优化将使用1日均线、2日均线、3日均线、5日均线……144日均线一共11种均线的长短组合进行搜索，寻找盈利能力最好的组合，从而确定最优的均线趋势策略形式。这里限定短期均线最大搜索到34日均线，长期均线从短期均线的后一根开始一直到144日均线结束（见表4-2）。整个设置都和上一节保持一致。

表4-2 所有组合下均线反转策略的年化收益率 (%)

长期均线＼短期均线	1	2	3	5	8	13	21	34
2	−4.92							
3	−6.16	−13.95						
5	−5.81	−13.16	10.82					
8	−12.58	−4.58	−1.20	−10.10				
13	−12.76	−16.67	−15.64	−17.06	−13.06			
21	−22.80	−22.29	−25.49	−21.63	−22.74	−22.97		
34	−21.45	−21.00	−22.09	−21.92	−21.45	−11.24	−8.46	
55	−36.32	−36.94	−44.20	−100	−44.53	−100	−12.24	−15.06
89	−43.18	−47.84	−32.20	−100	−100	−49.28	−49.17	−35.56
144	−29.87	−100	−50.15	−48.24	−41.00	−30.98	−26.35	−30.06

表4-2给出了所有的短期、长期均线组合下均线反转策略的年化收益率

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

情况。整个表中年化收益率为正的均线反转策略仅有 1 个，其年化收益率为 10.82%。剩下的组合年化收益率全部为负，甚至存在 5 种组合的年化收益率达到 −100%，即在交易过程中损失掉全部本金的情况。

从表 4-2 中的结果来看，网格搜索下的均线趋势策略相对而言并不是一个稳健的策略，绝大部分的搜索都落入收益率不佳的区间内，最优选择处呈现出一个陡峭的峰值。不过在这一节的简单框架之下，最后通过收益最优原则选择出来的短期均线为 3 日均线、长期均线为 5 日均线的均线反转策略，至少在样本上具有一定的收益表现。

在优化完成之后，再将短期均线为 3 日均线、长期均线为 5 日均线的均线反转策略应用到相同的样本数据之上进行回溯测试，从而得到回溯条件下相应的交易情况。图 4-8 展示了使用这一均线组合的均线反转策略进行交易的情况下，以 1 为初始值的净值走势情况。在 2011 年年底、2012 年年底、2013 年年底和 2014 年年底，交易策略下的净值分别变为 1.1172、1.5256、1.9528、1.5793，2015 年 6 月结束时净值达到 1.5878。从图 4-8 中也可以看到，自模拟交易开始起一直到 2013 年年底，该均线反转策略的净值水平都保持着上升的趋势，特别是在 2012 年前两个月和 2013 年下半年，净值增加比较明显。2014 年后净值在一定水平上维持了接近一年的时间，然后经过一个较大的波动下降到最后的 1.5878。基于整个交易过程来看，四年半的策略收益率为 58.78%，折算下来的年化收益率达到 10.82%。此外需要说明的是，图 4-8 和图 4-6 的纵轴比例尺不一样，而且后面的所有图形均是如此，使用图形进行对比的读者应注意纵轴坐标的具体刻度。

与上一节的均线趋势策略一样，从 2011 年 1 月 10 日收盘开始，均线反转策略的净值就再也没有下降到 1 以下的水平。在整个四年半的均线反转策略交易当中，净值的最大回撤率为 35.69%，最大回撤发生的起始点为 2014 年 4 月 1 日收盘的 2.1704，最低点则是 2015 年 4 月 27 日收盘的 1.3958，整个回撤的走势维持了一年以上。经过计算，整个交易过程中平均的年度最大回撤率为 16.36%。仅仅和上一节中的结果进行对比而言，均线反转策略在收益和风险两个方面都逊于均线趋势策略。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

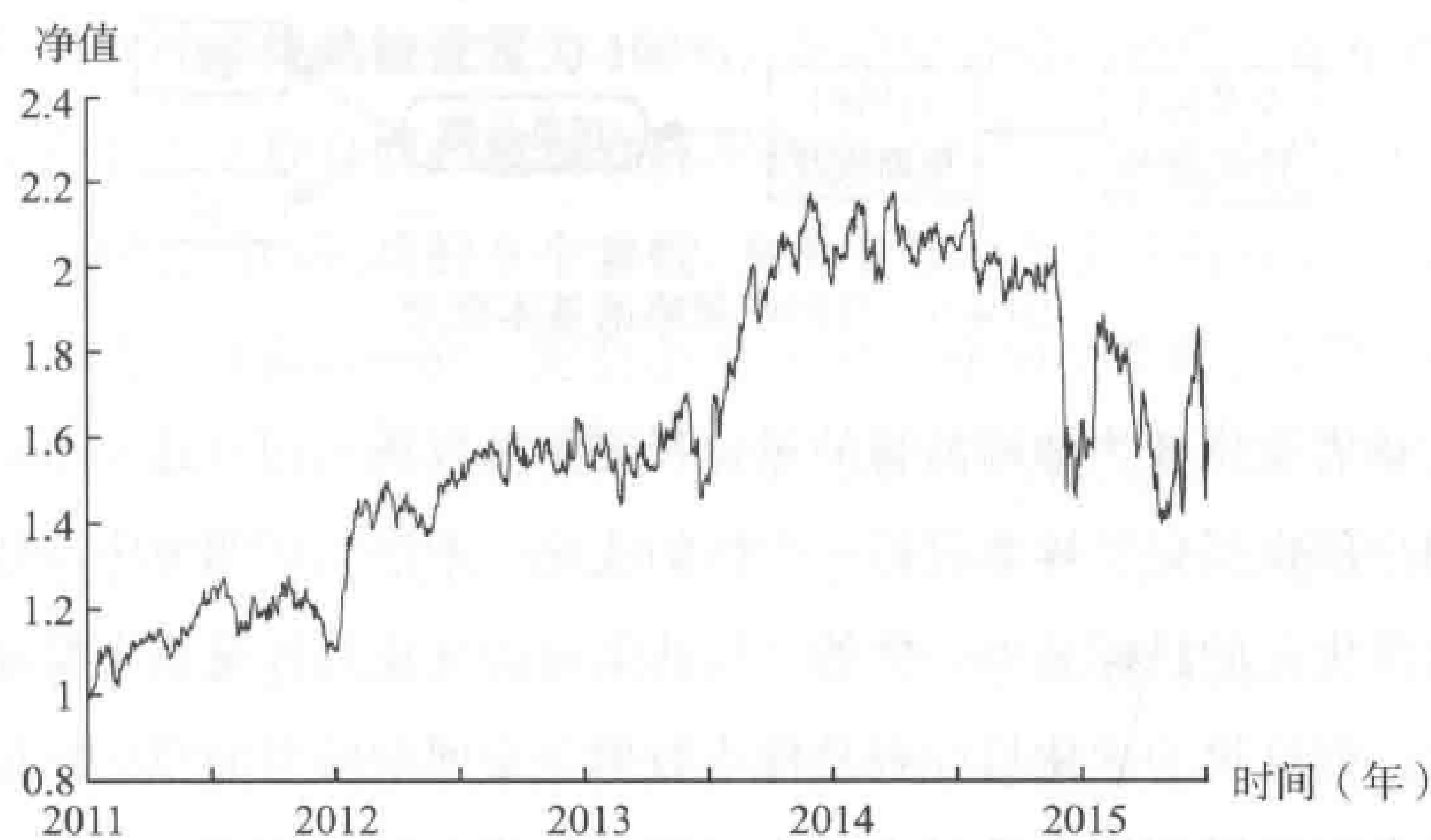


图 4-8 简单优化的均线反转策略净值走势

4.4 自回归策略的简单优化

自回归模型是计量经济学中最简单最基本的模型之一，常常用来刻画时间序列的性质以及做出相应的预测。这一节作者就通过自回归模型构建量化择时策略，并将其运用于现实数据当中，为基于收益预期的择时策略提供一个简单的案例。图 4-9 在图 4-4 的框架基础上，说明了一个自回归策略是如何运行的。首先假设自回归模型的形式如下：

$$r_t = \sum_{i=1}^{N} b_i r_{t-i}$$

式中，r_t 是交易资产在 t 时刻的收益率；b_i 是各阶的回归系数。

交易决策的时间点为 t 时刻头、$t-1$ 时刻末，此时 r_{t-1} 到 r_{t-N} 一共 N 日的日度收益率都是已知数据，通过计算上面的回归式就可以得到对 t 时刻收益率的预测值。当 t 时刻的预期收益率为正时，则看多 t 时刻的交易资产；当 t 时刻的预期收益率为负时，则看空 t 时刻的交易资产。分类过程中同样需要保证两个分类标准合并起来涵盖所有的可能性，因此当预测值为零时，固定将其分类为多或者固定将其分类为空即可。预测值精确为零的情况在实际研究中十分罕见，即使出现也影响不大。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

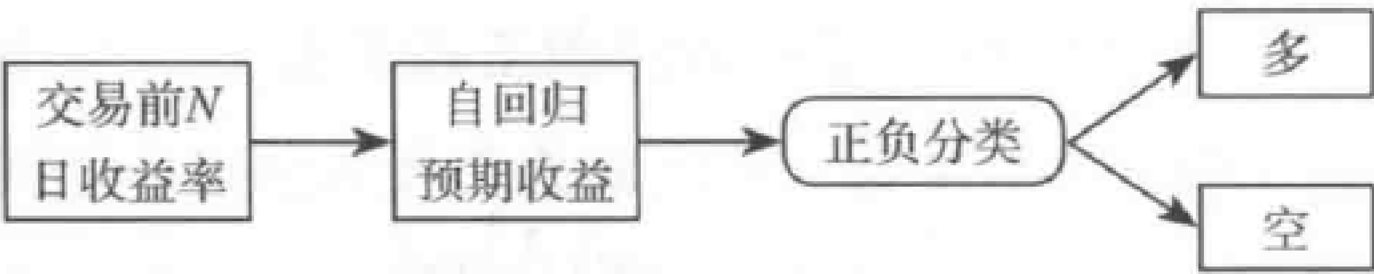


图 4-9　自回归策略的基本框架

为了给读者提供较为清晰易懂的量化择时策略以供入门，这一节同样在优化过程中使用全部模拟交易样本进行一次性的优化。不过与前两节中均线策略的优化相比，这里优化的目标是不一样的。均线策略的优化目标是整个交易期间内的收益率最大，而这里的优化目标则是样本数据上自回归模型的拟合度最高，从实践的角度来看就是模型拟合后对未来收益率的预测最好。前面已经提到过，对于这样的策略框架而言，由于存在着预测未来收益这样一个中间步骤，因此参数的最优化既可以通过自回归模型在样本中的拟合来完成，也可以通过使得自回归策略在样本中交易的收益最大化来完成。对于自回归模型这样一个相对而言比较简单的策略模型来说，其实两种优化方法都是可以使用的，相对而言后一种方法可能还更加贴近实际交易的意图一些。但是这一节中使用的优化方法确定为前面一种，即通过拟合自回归模型来获取系数值，这样做一是为了给读者呈现出不同优化方法的实际案例，二是因为在实际研究中如果使用更复杂一些的预测模型或手段，特别是拟合技术相对固定的一些模型如神经网络等，后一种优化方法的计算过于复杂，因此在实践中被应用的情况其实并不多见。

选用的交易资产也同样是沪深 300 股指期货主力合约，具体的样本数据包括从 2010 年 4 月 19 日到 2015 年 6 月 30 日的日收盘价。判断买卖从 2010 年 12 月 31 日收盘开始，也就是说在回溯测试样本中自回归策略一共模拟交易四年半的时间。2010 年 4 月 19 日到 2010 年 12 月 30 日的日收盘价数据作为交易前期计算 N 日日度收益率的备用数据，数据长度为 172 日，也就是说最高能支持 N 为 172 的自回归模型进行计算。在这一节的研究中，模型形式应该是在优化之前就确定好的，这里作者直接将自回归的阶数固定为 9，也就是将 N 取为 9（见表 4-3）。股

表 4-3　自回归模型拟合结果

滞后阶数	回归系数
1	0.01356
2	−0.05809
3	−0.01531
4	0.04056
5	0.03179
6	−0.03602
7	0.06777
8	0.05457
9	0.07993

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

指期货交易的杠杆率依然被设置为 100%，从而使策略的研究过程相对简单。

表 4-3 给出了 9 阶自回归模型在样本数据上的拟合估计结果，9 个回归系数实际上就是组成自回归策略的 9 个参数。9 个系数中为正的有 6 个，分别是 1 阶、4 阶、5 阶、7 阶、8 阶、9 阶，为负的有 3 个，分别是 2 阶、3 阶、6 阶。由于每一个滞后阶的日度收益率均值基本一致，因此可以直接用回归系数的大小来初步判断滞后收益率对未来收益率的影响程度。9 个系数的绝对值均在 0.01 以上，也就是说均具有一定的影响，但是滞后 1 阶的影响其实相对较小，系数估计值仅为 0.01356。滞后 2 阶的回归系数是 3 个负值中绝对值最大的，为 −0.05809，而正向影响最大的则是滞后 9 日的收益率，回归系数的估计值为 0.07993。不过，由于滞后的收益率之间也存在着相互关联关系，因此纯粹从系数大小来判断影响程度也是不可取的。如果要深究该问题，则需要进一步引入其他计量经济学的模型和检验方法，这里不做阐述。

在表 4-3 中，作者没有给出模型拟合下各个自回归系数的 t 统计量、p 值等数据来判断系数的显著性水平，也没有给出整体模型的 r^2 等数据用来判断模型的拟合程度等性质，原因在于量化交易策略的研究中虽然使用了计量经济学的模型，但是量化交易策略研发和计量经济学的研究目标以及研究思路都是有差别的。量化交易策略研发的根本目的是发现具有盈利能力的策略，而不是探究数据下的经济关系以及相应的解释和预测。当然，数据的检验与预测对于发现具备盈利能力的策略有着重要的推动作用，统计学和计量经济学的一些检验方法也能够在量化交易策略研发中起到指导性的作用。但是在本书的例子中，作者选择性地把一些计量经济学中重要的研究内容都忽略掉了，一是为了在大框架清晰的前提下尽量简化内容，二是希望读者能够通过作者这种舍弃的行为理解量化交易策略研发的一些本质特征。

在自回归模型拟合完成之后，便可通过确定了回归系数的模型在相同的样本下逐日预测下一日的收益率，进而判断交易方向，得到回溯条件下相应的交易情况。需要强调的是，由于策略逐日进行判断，同时我们又希望保证“股指期货交易的杠杆率为 100%”，因此约定在每个做空日之后都重新调整杠杆率为 100%。与前两节中均线策略做空期间不进行调整的情况是有所不同的，应特别注意。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

图 4-10 展示了使用这一系数确定的自回归策略进行模拟交易的情况下，以 1 为初始值的净值走势情况。在 2011 年年底、2012 年年底、2013 年年底、2014 年年底，交易策略下的净值分别变为 1.2181、1.1943、1.2553、1.2357，但是到了 2015 年 6 月交易结束，净值已经上升到 2.3943。从图 4-10 中也可以看到，在 2014 年 12 月中旬之前，该自回归策略的净值其实上升幅度较小，但是在 2014 年 12 月中旬之后，净值水平有了一个非常大的提升，从这个时间点开始到交易结束的 6 个半月间，净值的上升幅度达到 125% 左右。基于整个的交易过程来看，四年半的策略收益率为 139.43%，折算下来的年化收益率达到 21.41%。虽然收益情况不错，但是四年半的上升幅度其实主要贡献都来自最后半年多，这反映出该策略在样本期内缺乏一定的稳定性。

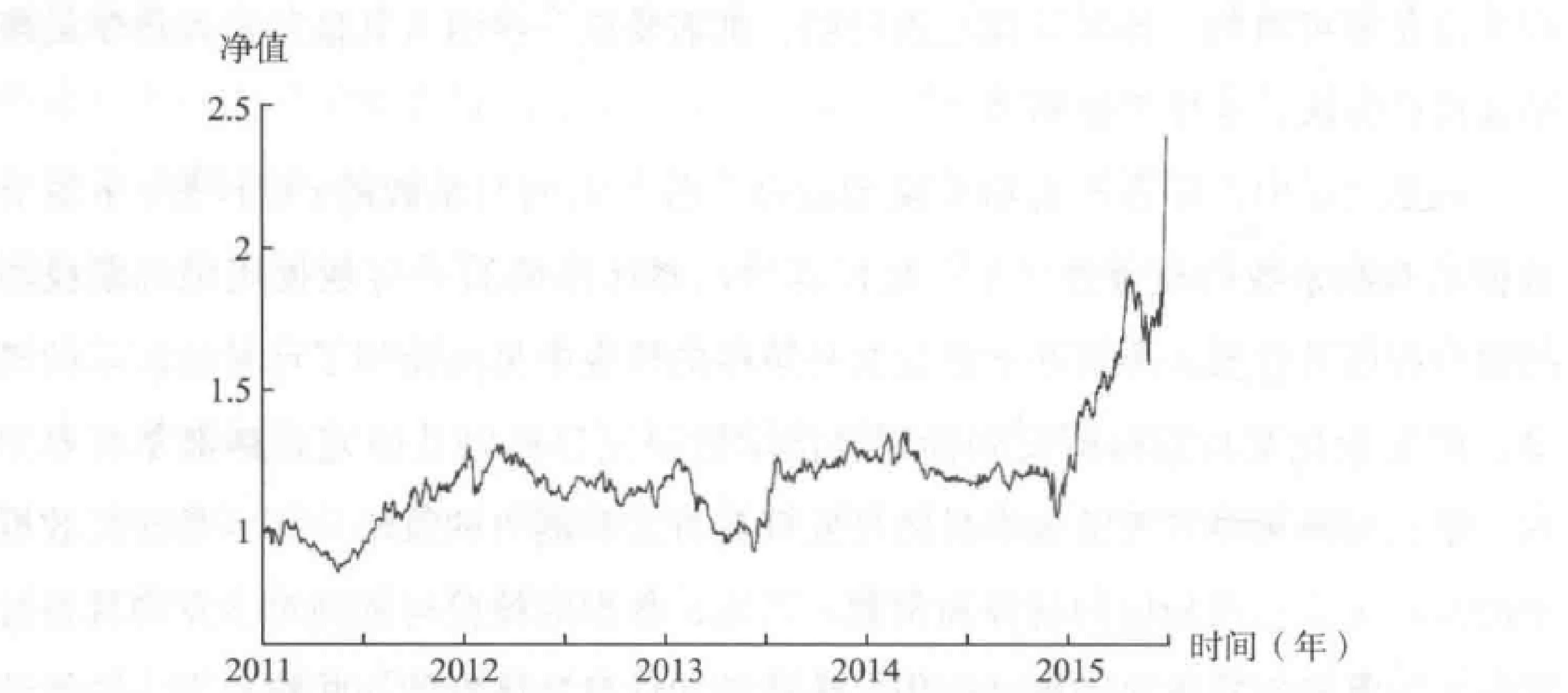
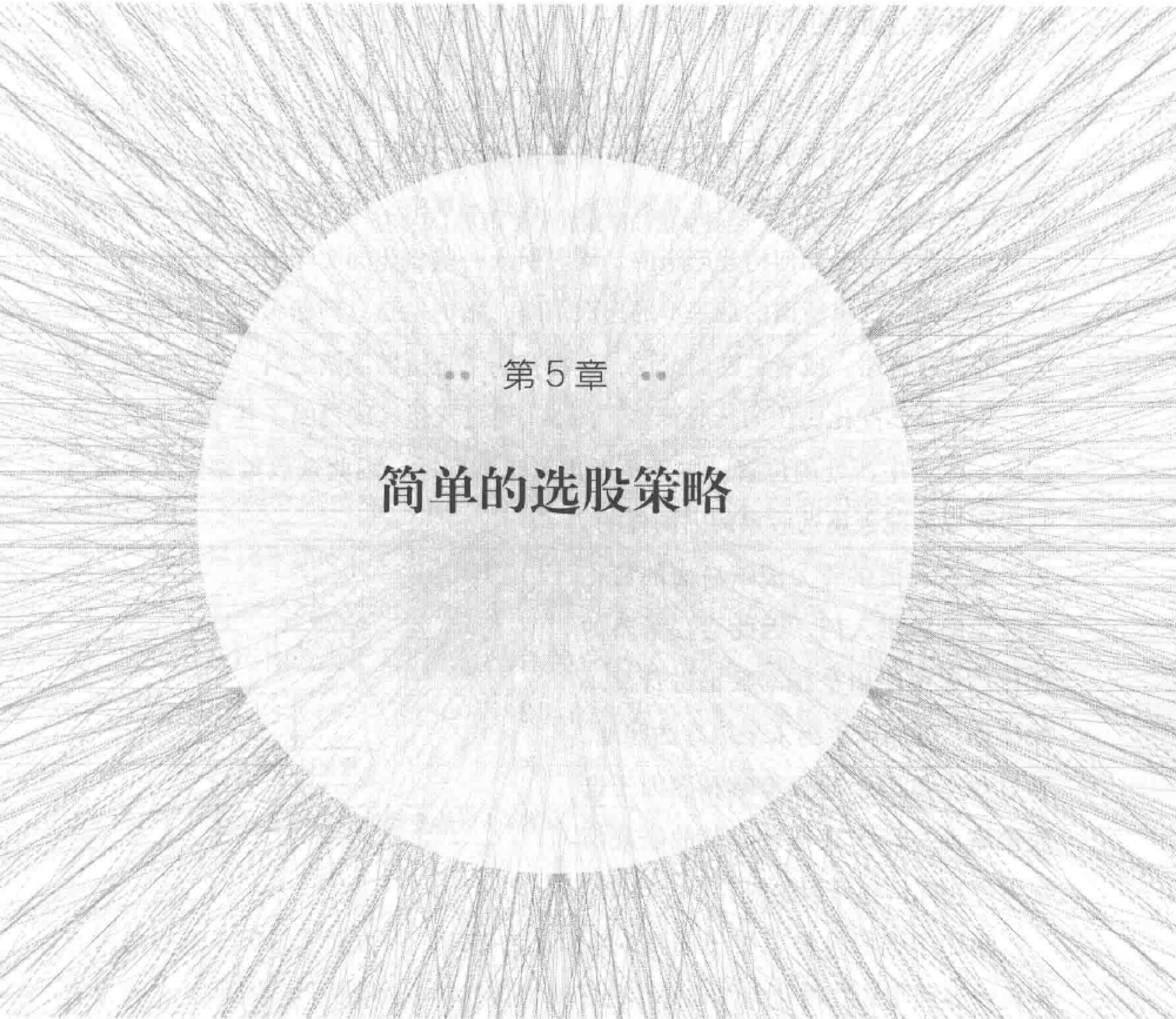


图 4-10 简单优化的自回归策略净值走势

在整个四年半的自回归策略交易中，净值的最大回撤率为 29.28%，这一次回撤的起始点是 2012 年 3 月 2 日收盘的 1.3048，最低点到 2013 年 6 月 18 日收盘的 0.9227，整个回撤的时间一共 310 个交易日。在这一次最大回撤之前，净值也曾在 2011 年 5 月 19 日收盘时下探到 0.8496。经过计算，整个交易过程中平均的年度最大回撤率为 18.53%。当然前面已有提到，整个策略的净值走势存在细节上的不稳定，因此仅靠年化收益率和平均年度回撤来判断策略是不够的，有必要结合净值的其他特征综合判断。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net



第5章

简单的选股策略

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

5.1 因子选股的基本框架

在上一章中，作者介绍了最基本的择时策略以及简单的一次性优化过程。作为与择时策略同样重要的选股策略，其在横截面上的作用与择时策略在时间轴上的作用是量化交易策略的两个重要方向。在这一章中，作者就向读者介绍最基本的量化选股策略是如何构建起来的，同时引入一些相应的实际案例用以说明。与上一章一样，这里所谓的最基本的选股策略，指的也是只判断买卖、不涉及仓位优化的选股策略，或者更贴切地说，是在挑选了合适的股票之后，只简单分配资金、不做具体配比优化的选股策略。图 5-1 通过灰色区域说明了基本的选股策略在研发流程中所处的位置。由于不考虑风险的问题，因此选股策略完全是通过判断收益情况来挑选股票的。同样的，为了给读者提供较为清晰易懂的量化选股策略以供入门，因此会在策略的优化过程中使用全样本数据进行简单的优化，进而产生出未来信息的问题。不过对于单因子选股策略和多因子选股策略而言，未来信息问题的表现情况不尽相同，在后面的内容中会具体说明。

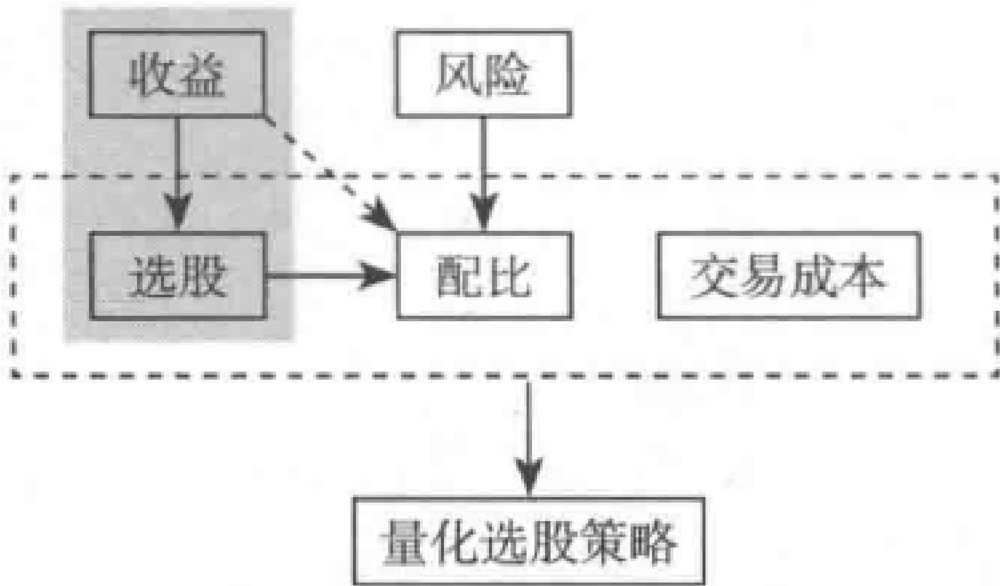


图 5-1　基本的选股策略在研发流程中的位置

单因子选股是整个因子选股策略框架的基础。一方面，单因子选股可以直接反映该因子的获利能力，如果单因子选股策略的收益情况达到要求，是可以直接用来进行策略交易的；另一方面，单因子选股的过程实际上与因子挑选密切相关，在大多数情况下，只有在单因子选股过程中能获得一定收益的因子，才有理由被纳入多因子选股策略的考量之中。其实相较于多因子选股策略的模型，单因子选股策略非常的简单直接，图 5-2 给出了一个最为粗略的单因子选股策略框架，在决定选股的时间点上，通过分类器来判断因子的取值是否符合一定条件，如果符合则将股票选入，在下一期持有。在这样的框架下，选股策略在表面上并没有涉及对后一期股票收益的预测，也没有使用一般意义下的最优化手段。不过就作者的看法，预测在分类时已经暗含其中，测试不同因子的过程也可以被看作

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

广义上的最优化。

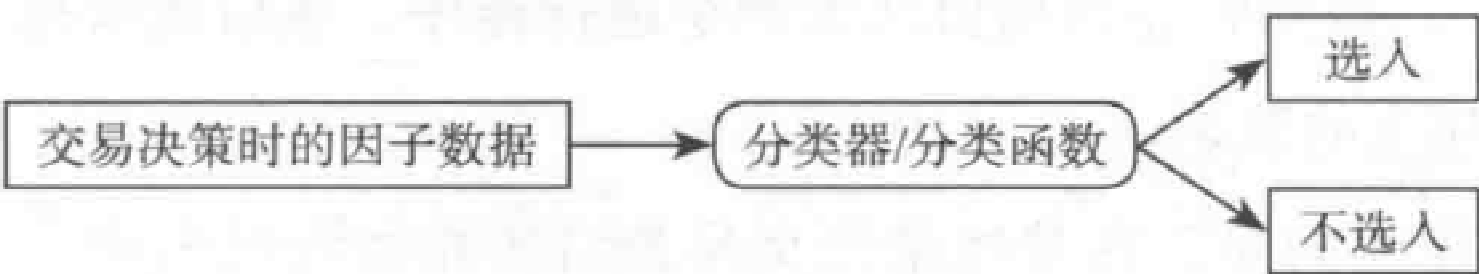


图 5-2 单因子选股策略的粗略框架

在实践中，因子选股已经是一个相对成熟而且大致框架较为固定的策略，同时存在一个业界从业人员都较为认同的策略体系，本书作为一个以介绍策略框架为主要目标的作品，也无意去变动或复杂化这样一个体系框架，因此在内容上基本介绍最简单、最受认可的策略构建方法。而操作单因子选股策略时，需要考虑到将单个因子融合进多因子策略的问题，因此在分类、判别的部分不宜过于复杂，甚至可以说存在一个较为固定的判别模式。简单而言，就是在交易决策时按照当前因子影响力的大小进行排序，然后选入排序前列的股票，排序选择顺序还是逆序则由下一期收益率与因子的具体关系决定，最后保证选取股票的收益率相对更高即可。图 5-3 给出了这样一个基本的单因子选股策略框架。

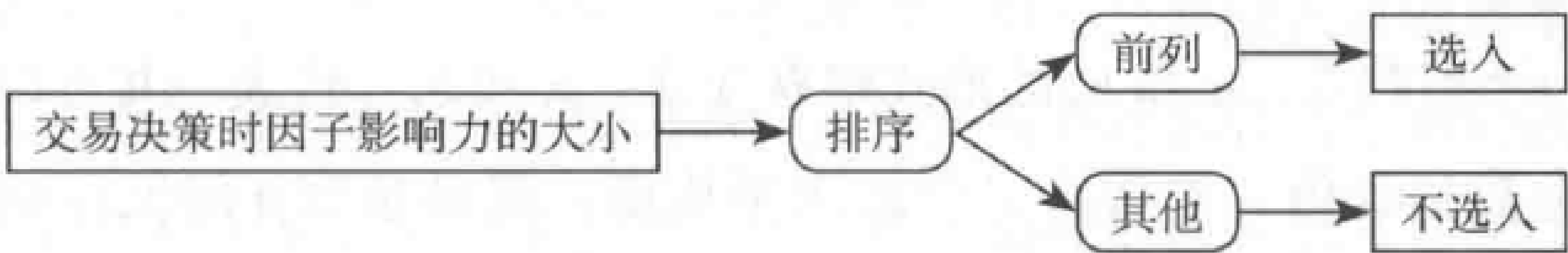


图 5-3 基本的单因子选股策略框架

多因子选股策略方面，本书也选用一个较为成熟的策略框架来进行介绍。首先假定已经选出了 N 个有效因子，分别用 f_1 到 f_N 标记，然后针对每一只股票，用这 N 个因子的值来预测下一期或者下一段时间的收益率。这里选用简单的线性回归来完成预测工作，如下式所示：

$$r_t = a + \sum_{n=1}^{N} b_n f_{n,t-1}$$

式中，r_t 是时刻 t 的股票收益率；$f_{n,\ t-1}$ 是 $t-1$ 时刻下第 n 个因子的大小；a 和 b_n 是回归式中的系数。

进行交易决策的时间点为 t 时刻初、$t-1$ 时刻末，因此回归式左边为预测值，

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

回归式右边的所有成分则都是决策点下的已知信息。在预测出每一只股票在时刻 t 的收益率之后，按照收益预期值从大到小进行排序，然后选取排在前列的股票作为当前可以建仓的股票。

需要特别说明的是，在某些量化交易策略的相关资料当中，会把对于不同股票而言取值一致的回归系数 b_n 称为风险因子，而将具体的股票特征值 $f_{n,\ t-1}$ 称为各只股票在因子上的溢价。这主要是因为学术界在套利定价理论等研究的基础上，形成了一种约定俗成的叫法，其中风险因子对于所有资产应该保持一致，而因子溢价则各有不同。不过在量化选股策略中，对比本节所使用的称谓，这种叫法以及其他一些叫法并不是非常直观，因此不予以使用。如果读者在阅读其他资料时碰到不一样的名称，只需对号入座弄清准确含义即可。[⊖]

a 和 b_n 等参数的优化和拟合，书中使用的是法玛等人给出的一种线性回归估计方法。如果可以获得 T 个时间段的因子数据以及相应的下一期股票收益率数据，那么对于上面的线性回归式而言，一共可以进行 T 次估计，表示如下：

$$r_t = a_t + \sum_{n=1}^{N} b_{n,t} f_{n,t-1}, \quad t=1, \cdots, T$$

相比起上一个回归式，a 和 b_n 的形式略有变动。a_t 和 $b_{n,\ t}$ 代表一共可以得到 T 组 a 和 b_n 的参数估计值。然后将 T 个 a_t 求平均值，就是参数 a 的拟合结果；分别将 T 个 $b_{n,\ t}$ 求平均，则得到 N 个 b_n 各自的拟合结果。

图 5-4 展示了基本的多因子选股策略的基本框架，和单因子选股策略不同的地方在于，单因子选股直接针对因子进行排序来选择股票，而多因子选股需要先基于因子预测股票的未来收益，然后对预期收益进行排序进而选择下一期持有的股票。因此，多因子选股策略中很明确地包含了预测的成分，以及通过估计回归模型来完成的最优化手段。如果像第 4 章一样仅介绍简单的选股策略，那么从表面上来看，多因子选股策略的研究过程中由于存在回归模型的一次性估计，因此包含未来信息的成分，而单因子选股策略的研究过程中则不存在未来信息的问题。但是正如前面所言，不断地尝试不同的因子、最后挑选出合适因子的过程，

⊖ 例如在 BARRA 的资料中，又会将回归系数称为因子收益，将具体的股票特征值称为因子的风险暴露。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

实际上也是在进行最优化的筛选，这其中暗含的未来信息问题，需要研究人员加以注意。

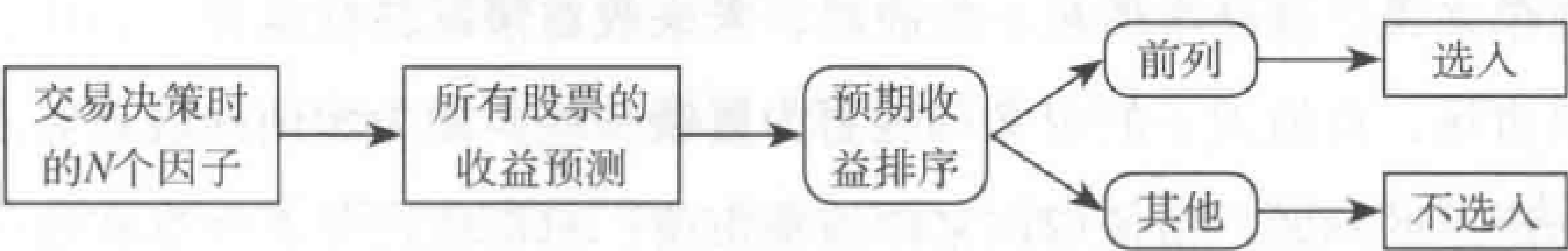


图 5-4 基本的多因子选股策略框架

在本节开始处作者提到过，单因子选股模型一方面可以直接反映该因子的获利能力，另一方面可以过渡到多因子选股策略的准备工作——因子挑选之上。不过单因子选股模型在这两个方面发挥作用的时候，策略研发人员的关注点和需要处理的细节问题都有细微的差别，图 5-5 给出了一个简单的例子用来说明。当研究人员使用单因子模型直接构建交易策略进行交易时，主要关注的是被选入股票的收益情况，如图中曲线的加粗部分所示。只要这一部分的未来收益能够达到标准即可，至于因子大小排序在后面的股票，不管是展现出如图中的非线性趋势，还是有明确的线性趋势甚至是没有一个显著的走势，在重要性上要远逊于选入股票的收益情况，至多影响到策略的进一步改进。但是当研究人员使用单因子模型来挑选有效因子从而构建多因子模型时，由于涉及具体的收益预测问题，单个因子在多因子预测中需要合并起来使用，因此整个曲线的性状都是要加以考虑的问题。而本书中所述的多因子策略方法基于线性回归模型，因此预期收益和因子大小之间符合线性关系是最优状态，当两者的关系呈现出如图 5-5 中所示的非线性关系时，在线性回归步骤之前对因子或收益数据做合适的预处理，例如法玛所使用过的对数化，就是一个很有必要的过程。

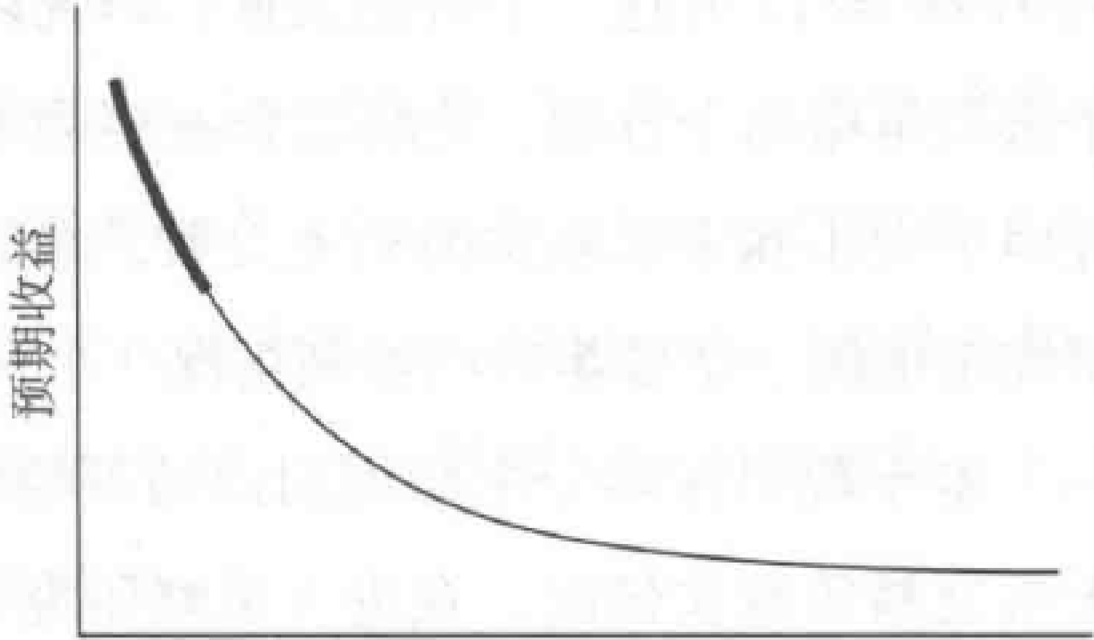


图 5-5 单个因子与预期收益关系的简单示例

5.2 市值因子

市值因子一直是业界和学术界都比较认同的一个选股因子，法玛标志性的三

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

因子模型当中，市值因子就是其中之一。法玛在他的实证研究中发现，就美国市场上的股票而言，随着股票市值的上升，其未来收益率存在一个显著的逐渐下降趋势。简单来说，就是市值越小的股票，未来收益情况往往越好。当然，对于中国的股票市场，市值因子的定义可能更为复杂一些，因为国内的股票更多地涉及流通与非流通的问题，市值应该采用流通市值、总市值还是其他形式进行计算是一个存在选择性的问题，好在实际的策略研究发现市值因子定义上的细微变化并不显著影响到其盈利能力的存在与否。而且实际上，市值因子对于目前的中国股票市场而言，可以说是接近公认的最为有效的选股因子。虽然就在本书所涉及的样本期的末尾，市值因子在一段时间内完全失效，但是从后续的市场运行情况来看，市值依然是中国股票市场上一个不能被忽略的重要因子。

为了研究这个选股因子，作者获取了自 1991 年 1 月开始到 2015 年 6 月底结束的国内股票月度数据。所选股票的范围涵盖在深圳证券交易所进行交易的所有 A 股，包括主板、中小企业板和创业板三个板块，以及在上海证券交易所进行交易的所有 A 股。月度数据则包括月末收盘时股票的复权后月度收益率，以及股票的月度流通市值。开始挑选股票进行买卖的时间从 2010 年 12 月的收盘开始，也就是说在回溯测试样本中将一共交易四年半的时间，和上一章中的简单择时策略的交易时间保持一致。1991 年 1 月到 2010 年 12 月的所有月度流通市值数据则作为开始买卖股票前的备用数据，不过对于单因子的简单判断来说，只需要用到 2010 年 12 月这一个月的数据。研究内容上将同时兼顾因子有效性测试和单因子选股策略两个方面，先将股票按照因子从小到大排序分成若干个等份，从而研究因子大小和未来收益的相关关系，然后选取排列中的第一等份建立仓位，进而研究市值因子选股策略的盈利情况。

鉴于股票在整个生命周期内的不稳定性，也就是可能出现的停牌、涨跌停、无法交易等现实情况，在进入策略研究阶段之前，需要对相应的数据进行处理，尽量贴近真实环境下的交易情况。对某一些在研究过程中无法完全处理的差异，也应该做到心中有数，如有可能尽量在结果中进行估计和描述。而在整个回溯过程中，模拟交易也设定为只进行买入和卖出操作，不进行卖空的操作。这样设定的原因在于中国的股票市场本身处在一个高速发展的阶段，相关的交易手段并不

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

完善而且变化较快，直到本书写作的时间也只有少部分股票可供做空，因此回溯测试中完全体现出现实的交易环境相对而言比较麻烦，做出不卖空的交易限制则可以有效地简化整个量化选股策略，方便读者更好地掌握相关内容。同时在本章的基本选股策略框架下，全部资金都用来持有股票。

回溯测试的过程其实比较简单，在第一个选股和交易的时间点，即 2010 年 12 月收盘时，将样本中所有数据完备且可以交易的股票当前的流通市值从低到高进行排序，然后按顺序等分为若干个部分，每一等份中的股票又按照等权重进行组合。也就是说，把每一个等份看作一个股票组合，则该组合中按照收盘价买入的每一只股票的资金量都相同。然后在 2011 年 1 月收盘时进行相同的排序和等分过程，形成新的股票组合分组。之后每一个月收盘时都进行相同的操作，直到交易样本结束即 2015 年 6 月底为止。

在 54 个月份之间，排序第一的等份就对应着一个长度为 54 的收益率序列，假设在 2010 年 12 月收盘的时候资金量为 1，那么对于“在下个月持有当月排序第一的等份中对应的股票”这样一个策略，就可以通过该收益率序列计算出一个净值走势，这也就是单因子选股策略的净值走势。需要说明的是，除了交易成本在这一章的简单研究过程中暂时不考虑以外，这里的策略回溯过程还将引入另一个与实际交易环境不符的刻画：如果在一个收盘交易点判定了新的选入股票之后，原来某个正在持仓的股票不在新的选入股票之列，但是此时由于该股票因为停牌等情况无法卖出，那么在回溯环境下假设该股票同样退出持仓，且该月的收益率为零。这样的假设虽然脱离现实，但是会使得回溯过程更加清晰，在实际交易中这种情况会涉及仓位的分配调整，是一个需要研究人员自行做出处置的疑难问题。

对于排序不同的等份而言，均可以像处理排序第一的等份一样，将其看作一个逐月变化的股票组合并计算出净值变化的情况，再折算成月度的收益率，代表的就是这一排序的平均月度收益，如下式所示：

$$\overline{r}_n = \sqrt[54]{\prod_{t=1}^{54}(1+r_{n,t})} - 1$$

式中，$r_{n,t}$ 是排序第 n 的等份组合在第 t 个月的收益率；$\overline{r}_n$ 就是排序第 n 的等份在样本区间内的平均月度收益率。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

如果因子大小和未来收益有显著的关系，那么$\overline{r}_n$应该随着 n 的变化具有显著的趋势。当然，这种趋势也会随着 n 的具体取值而变化，如果 n 太小，不同等份间变动的表现能力会因为太多的股票参与平均而减弱，如果 n 太大，趋势又可能会受到股票收益中噪声的影响而不明晰。

在本书中，作者将所有股票按照因子大小排序划分为 3 等份、10 等份、30 等份共三种情况来加以研究。

5.2.1 市值因子选股 3 等分

图 5-6 展示了将所有股票按照流通市值大小划分为 3 个等份，每一等份中的股票又按照等权重进行组合时，3 个等份各自的平均月度收益，分别为 2.62%、1.46% 和 0.94%。由于同期的股市从整体上来说是上涨的，因此三个值均为正值是一个合理的结果。第 1 等份和第 2 等份之间的月度收益率相差 1.16%，第 2 等份和第 3 等份之间则相差 0.52%，表现出了非常明显的下降趋势。当然，由于等份数量较少，只以此来判定因子大小和未来收益间的逆向关系是不够的。

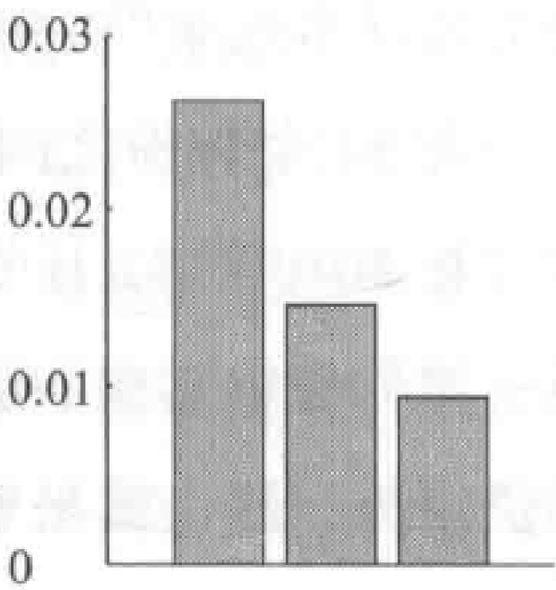


图 5-6 市值因子选股 3 等分的收益情况

在确定了逐月的排序和 3 等分之后，就可以顺理成章地得到第 1 等份股票组合的逐月变动情况，也就是市值因子选股策略的交易详情。图 5-7 中的实线展示了使用市值因子选股策略进行交易的情况下，以 1 为初始值的策略净值走势，虚线则是同期沪深 300 指数的走势，用以比较。可以发现，该选股策略的净值在 2014 年之前没有明显的走势，只是稍有下探然后又回复到初始值左右，2014 年后显示出了一个比较明显的上升趋势，到 2015 年 6 月底净值达到 4.0429。从整个交易过程来看，四年半的策略收益率为 304.29%，折算下来的年化收益率为 36.40%。对比而言，同期四年半的沪深 300 指数的收益率则仅为 42.99%，经过折算，指数的年化收益率为 8.27%。通过比较收益率情况可以发现，基于市值因子的选股策略在盈利能力上是优于市场指数的。此外，为简便起见，后面的研究内容中将不再说明沪深 300 指数的具体收益率，仅在图中以虚线标出作为对比，请读者注意。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

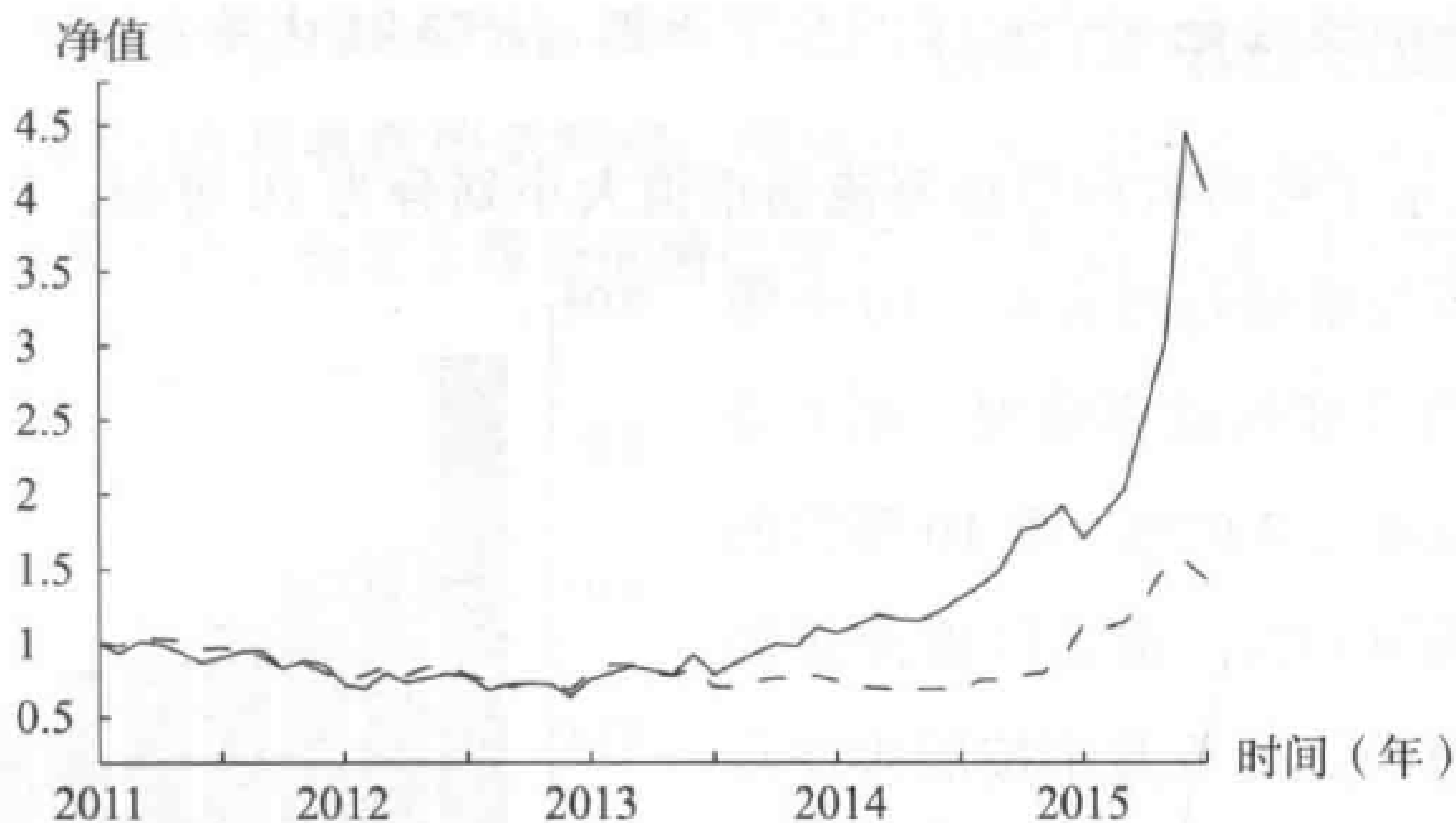


图 5-7 市值因子选股前 1/3 的股票组合净值走势

图 5-8 则是逐月地将市值因子排序前 1/3 的股票组合收益率减去沪深 300 指数的收益率后，所得到的收益率序列形成的净值走势⊖。净值从 2010 年 12 月收盘的 1 开始，在 2015 年 6 月底达到 2.5842，四年半的收益率为 158.42%，折算下来的年化收益率为 23.49%。四年半之间的最大回撤率为 40.38%，回撤的开始点为 2014 年 9 月收盘的 2.2646，结束点则是 2014 年 12 月收盘的 1.3503，实际上四年半中的主要回撤就是最大回撤下的这三个月。由于最大回撤的特殊性，整个交易过程中平均的年度最大回撤率这一指标在一定程度上失去了评判的意义，因此不再具体给出。

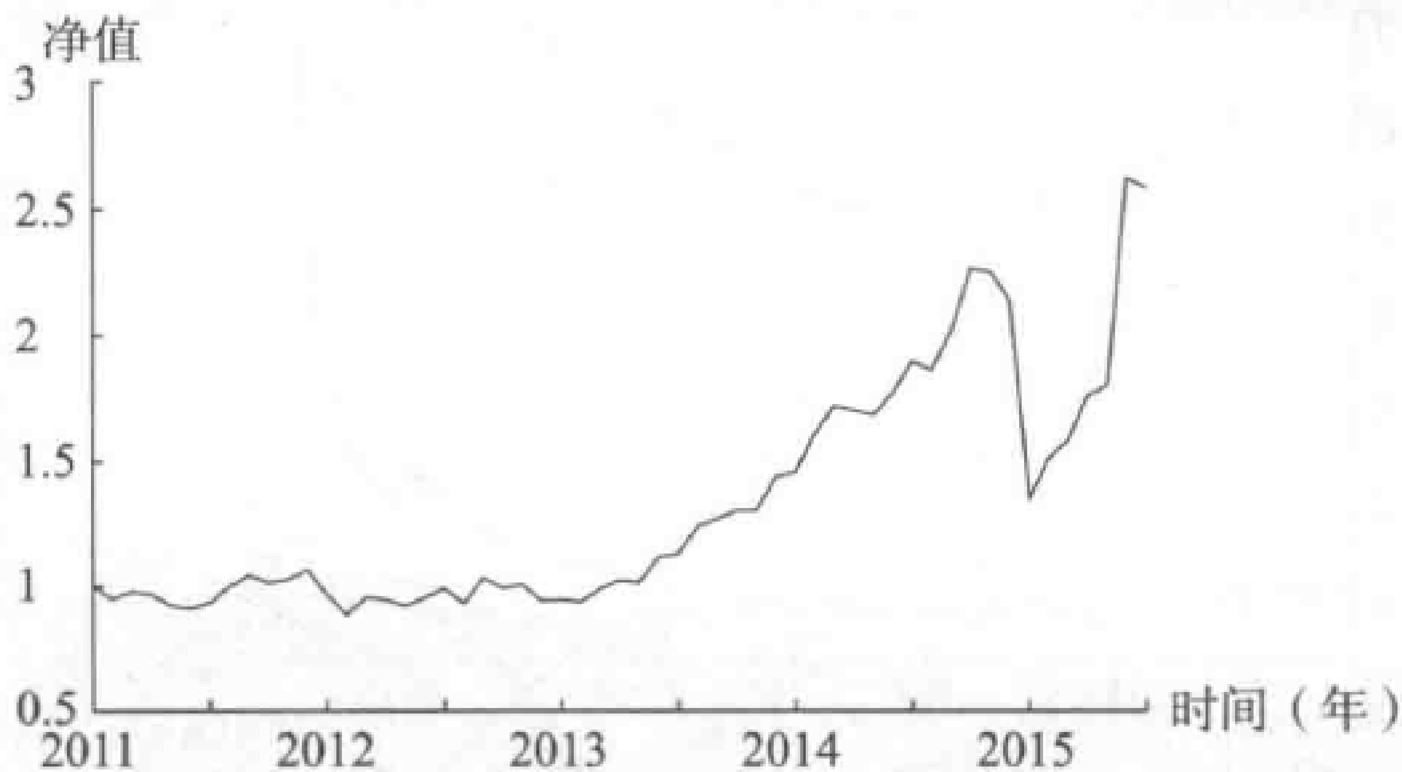


图 5-8 市值因子选股前 1/3 的股票组合与沪深 300 收益之差的净值走势

⊖ 虽然这样处理并不具备很强的现实意义，与实际交易中的中性化操作有或多或少的差距。但是通过这样的简单相减并假想为一个新的交易结果，可以清晰地观察到策略相对于指数的表现情况，是作者比较偏好的一种策略评价方式。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

5.2.2 市值因子选股 10 等分

图 5-9 展示了将所有股票按照流通市值大小划分为 10 等份，每一等份中的股票又按照等权重进行组合时，10 个等份各自的平均月度收益率情况。第 1 等份的月度收益率为 3.62%、第 10 等份的月度收益率为 0.71%，而且值得注意的是，10 个等份的月度收益率按照排序存在着非常规律的下降顺序。这表明在 10 等分的情况下，市值因子的有效性得到了比较好的体现。

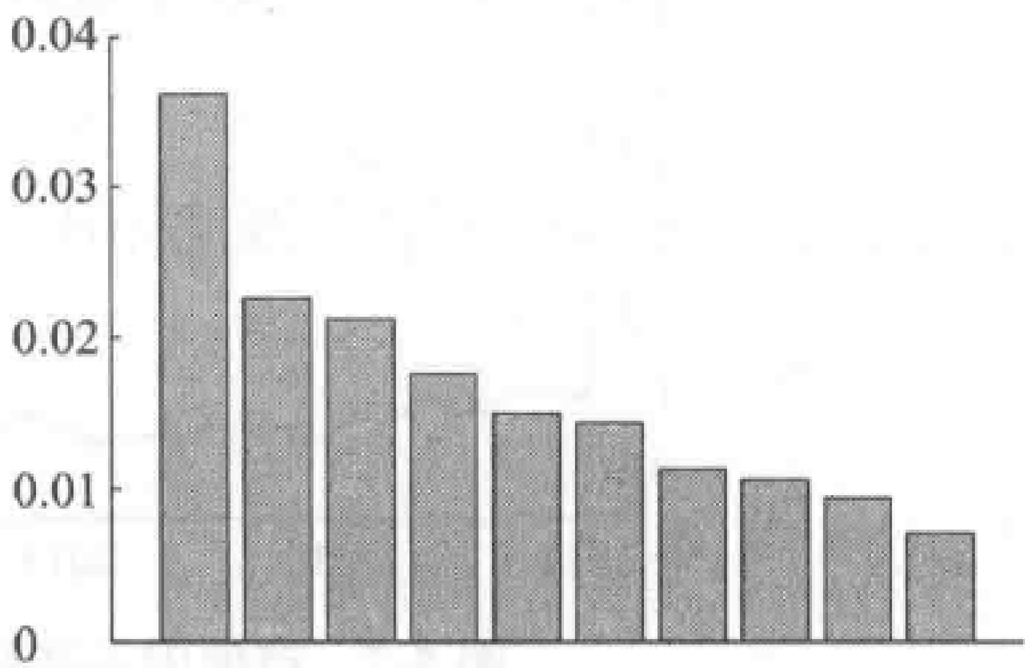


图 5-9 市值因子选股 10 等分的收益情况

在确定了逐月的排序和 10 等分之后，也就得到了通过市值因子选出的第 1 等份股票组合的逐月变动详情。图 5-10 中的实线展示了使用市值因子选股策略进行交易时，以 1 为初始值的净值走势情况，虚线则是用来对比的沪深 300 指数。相较于图 5-7 而言，净值走势在 2015 年的半年之间上升的幅度更加明显，这也使得 2015 年之前的走势显得越发平缓。策略净值在 2015 年 6 月收盘达到 6.8180，四年半之间策略的收益率一共为 581.80%，折算下来的年化收益率为 53.20%。

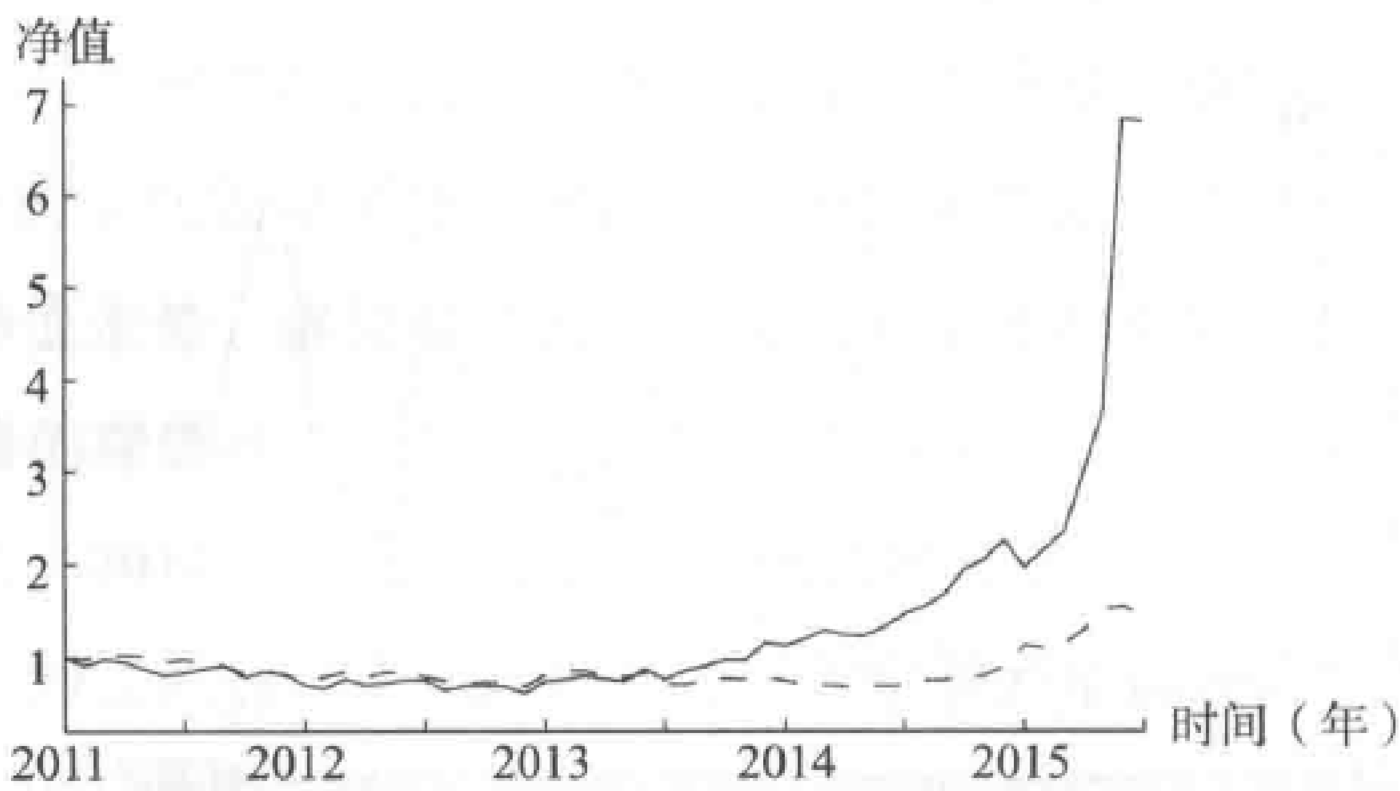


图 5-10 市值因子选股前 1/10 的股票组合净值走势

图 5-11 则是逐月地将市值因子排序前 1/10 的股票组合收益率减去沪深 300 指数的收益率后，所得到的收益率序列形成的净值走势。净值从 2010 年 12 月收盘的 1 开始，在 2015 年 6 月收盘达到 4.3541，四年半的收益率为 335.41%，折

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

算下来的年化收益率为 38.67%。四年半之间的最大回撤率为 39.91%，回撤的开始点为 2014 年 10 月收盘的 2.5743，结束点则是 2014 年 12 月收盘的 1.5469，回撤期长度为两个月，较之 3 等分的情况少了一个月的时间，即 2014 年 10 月。

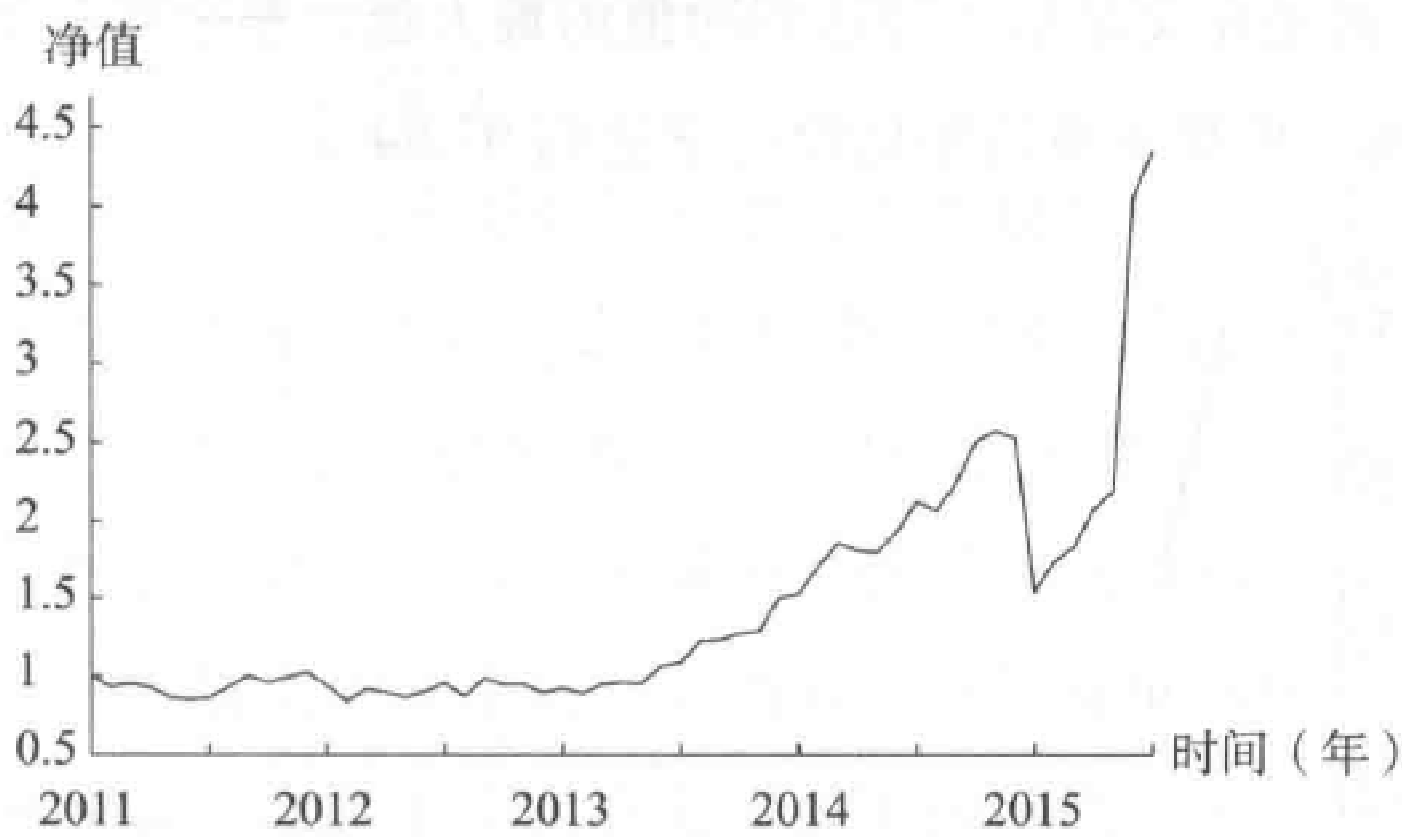


图 5-11 市值因子选股前 1/10 的股票组合与沪深 300 收益之差的净值走势

5.2.3 市值因子选股 30 等分

图 5-12 展示了将所有股票按照流通市值大小划分为 30 等份，每一等份中的股票又按照等权重进行组合时，30 个等份各自的平均月度收益率。最高收益率是第 1 等份的 5.67%，第 2 等份的收益率就下降至 2.32%，降幅在图中表现得非常明显，而后面的 29 个等份则基本表现出一定的下降趋势，第 30 等份的收益率情况为 0.75%。

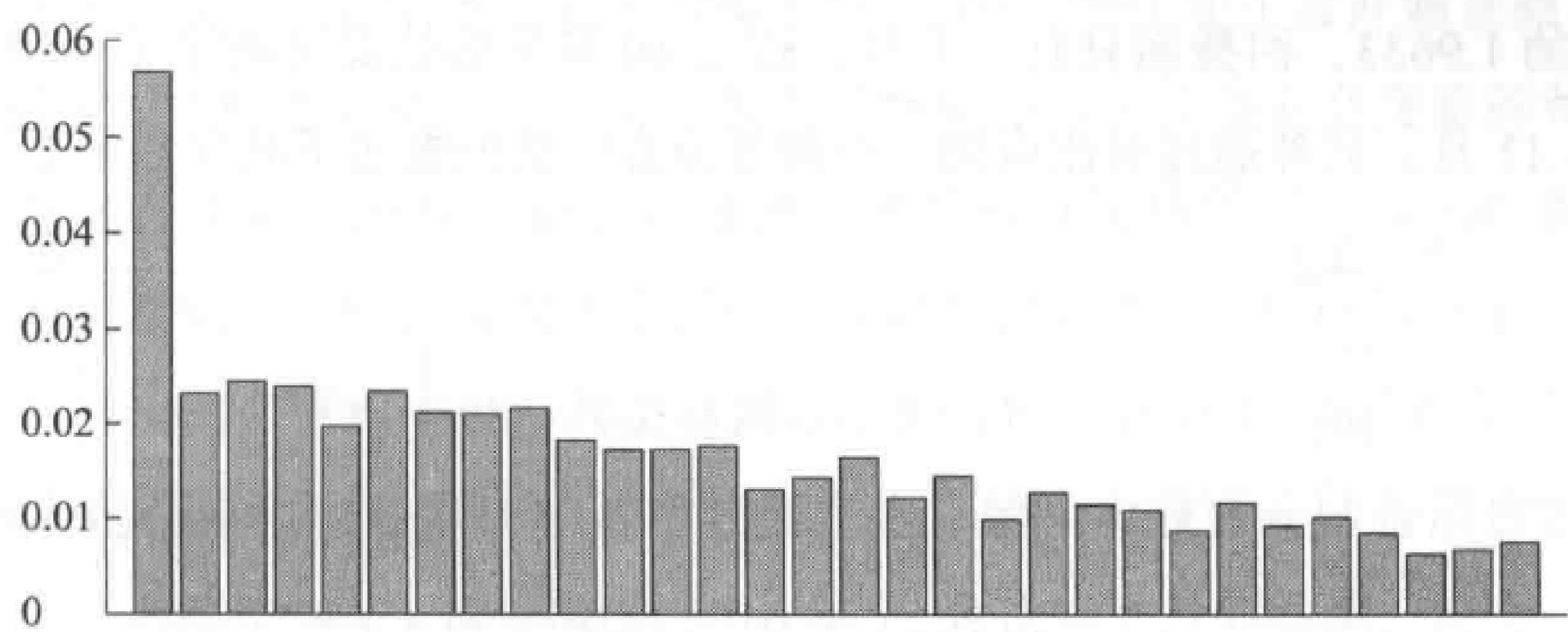


图 5-12 市值因子选股 30 等分的收益情况

在确定了逐月的排序和 30 等分之后，也就得到了通过市值因子选出的第 1 等份股票组合的逐月变动详情。图 5-13 中的实线展示了 30 等分的情况下，以 1

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

为初始值的市值因子选股策略净值走势，虚线则是用来进行比较的沪深 300 指数。该图的走势情况与图 5-10 非常相似，但是整个交易过程的净值增幅更加突出，2015 年 6 月收盘时的策略净值为 19.6552，而且末尾没有如图 5-7 和图 5-9 那样略有下降，而是在交易结束时达到净值的最大值。整个四年半期间的策略收益率为 1865.52%，折算下来的年化收益率达到 93.84%。

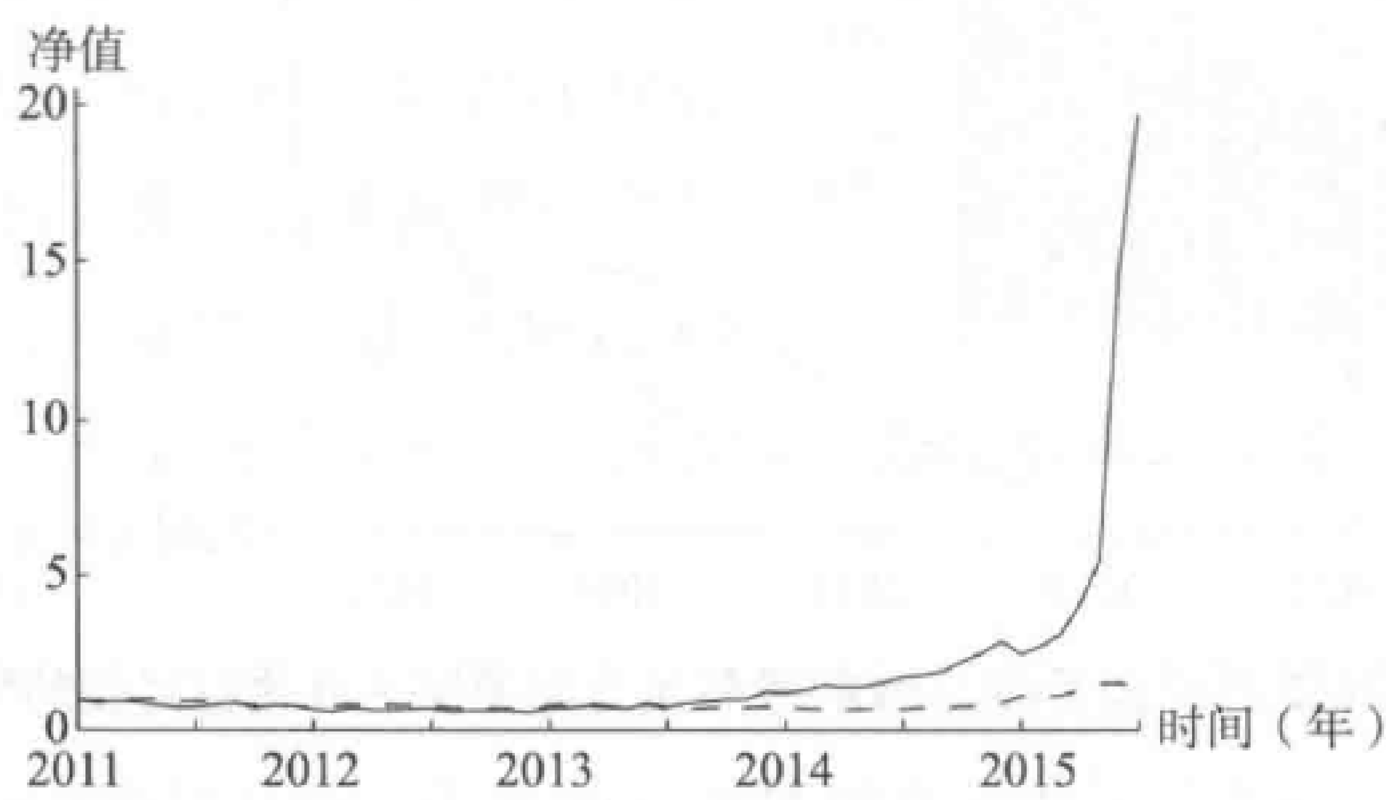


图 5-13 市值因子选股前 1/30 的股票组合净值走势

图 5-14 则是逐月地将市值因子排序前 1/30 的股票组合收益率减去沪深 300 指数的收益率后，所得到的收益率序列形成的净值走势。净值从 2010 年 12 月收盘的 1 开始，在 2015 年 6 月收盘时达到 12.5894，四年半的收益率为 1 158.94%，折算下来的年化收益率为 75.57%。四年半之间的最大回撤率为 39.89%，回撤的开始点为 2014 年 11 月收盘的 3.2663，结束点则是 2014 年 12 月收盘的 1.9633，回撤期只有一个月，较之 10 等分的情况又减少了一个月，即 2014 年 11 月。另外通过对比可知，三种等分方法的回撤水平其实相差都不大。

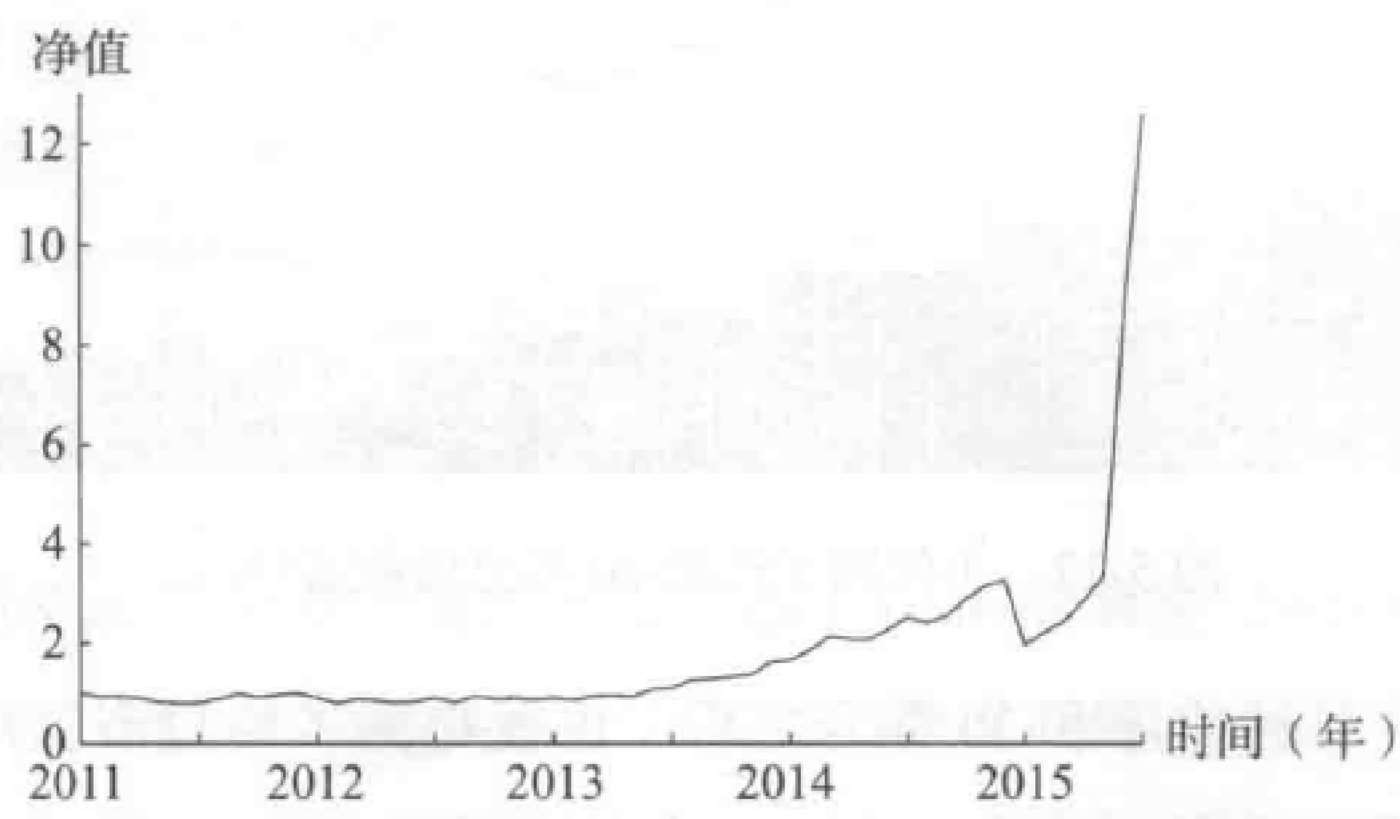


图 5-14 市值因子选股前 1/30 的股票组合与沪深 300 收益之差的净值走势

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

5.3 反转因子

反转因子是另外一个比较重要的选股因子，在实际的金融市场上，股票往往对当前的信息存在过度反应的现象，因此在短期内具有回复到正常状态的动力，从股价上看就是一种反转的行为。

一般情况下，如果研究使用的是月度的股票数据，那么常常将上一个月的收益率作为反转因子使用，从而描述短期内资产价格对信息的过分反应。在阿斯内斯等人的四因子模型中，普通的动量因子一般定义为一年前到上上个月的收益率，略过上个月的收益情况，也正是因为这种反转现象的存在。这种过度反应导致的反转，会使得股票的未来收益率随着上月收益率的增加而呈现出逐渐下降的趋势，简单来说，就是上个月收益率越低的股票，下个月的收益情况往往越好。

从后面的实际数据分析中可以发现，对于中国的股票市场，上个月收益反转因子同样是一个存在盈利能力的选股因子。但是值得注意的是，其在样本期的末尾也经历了一段失效时间，并且与市值因子的失效时间几乎重合。因此在那几个月中，使用因子选股策略的基金普遍产生亏损并不简单的市值因子失效的问题，市场特征发生变化导致若干个有效因子同时出现问题才是深层次的原因。

为了研究这个选股因子，作者获取的同样是自 1991 年 1 月开始直到 2015 年 6 月结束的国内股票月度数据。所选股票的范围涵盖在深圳证券交易所和上海证券交易所进行交易的所有 A 股，月度数据则是所有这些股票在月末收盘时的复权后月度收益率。开始挑选股票进行买卖的时间也从 2010 年 12 月收盘开始，也就是说在回溯测试样本中将一共交易四年半的时间。1991 年 1 月收盘到 2010 年 12 月收盘的所有复权后月度收益率数据作为开始买卖股票前的备用数据，不过对于单因子的简单判断只需要用到 2010 年 12 月这一个月的数据。研究内容也跟上一节一样，同时兼顾因子有效性测试和单因子选股策略两个方面，先将股票按照上个月收益反转因子排序分成若干个等份，从而研究因子大小和未来收益的相关关系，然后再选取排序中的第一等份建立仓位，进而研究反转因子选股策略的

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

收益情况。整个交易过程也设定为只进行买入和卖出操作，不进行卖空操作，资金全部持有股票。

需要特别指出的是，上个月收益反转因子和市值因子都有一个比较好的特性，那就是数据在金融资产进行交易时可以实时获取，该特性源于这两个因子同属于市场数据。当量化交易策略研发人员研究的不是市场数据而是某些财务数据时，这些数据可能会由于财务报表披露的延时而无法实时得到，交易时间点下数据库所显示的数据往往在实际交易中是无法获得的，也就会产生在第 3 章中提到过的未来信息的问题。上个月收益反转因子和市值因子都不存在这样的问题，回溯测试中按照数据库所显示的数据直接进行研究即可。这也是作者选择这两个因子作为案例的原因之一。

与上一节相同，作者将所有股票按照因子大小排序划分为 3 等份、10 等份、30 等份共三种情况来加以研究。

5.3.1 反转因子选股 3 等分

图 5-15 展示了将所有股票按照上月收益率大小划分为 3 等份，每一等份中的股票又按照等权重进行组合时，3 个等份各自的平均月度收益率情况。具体的收益率大小分别为 1.95%、1.71% 和 0.84%。第 1 等份和第 2 等份之间的月度收益率相差较小为 0.24%，而第 2 等份和第 3 等份之间的差距则更大一些，为 0.87%，约等于第 3 等份本身的收益率。划分为 3 个等份的情况下，下降趋势是比较明显的。

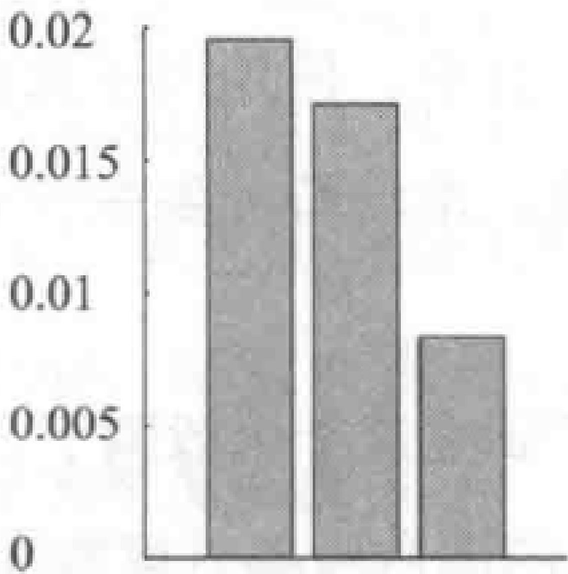


图 5-15 反转因子选股 3 等分的收益情况

确定了逐月的排序和 3 等分之后，就可以得到第 1 等份股票组合的逐月变动情况，也就是反转因子选股策略的交易详情。图 5-16 中的实线展示了使用反转因子选股策略进行交易的情况下，以 1 为初始值的策略净值走势，虚线则是用来对比的同期沪深 300 指数走势。策略净值在 2015 年 6 月底达到 2.8431，四年半的策略收益率为 184.31%，折算下来的年化收益率为 26.14%。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

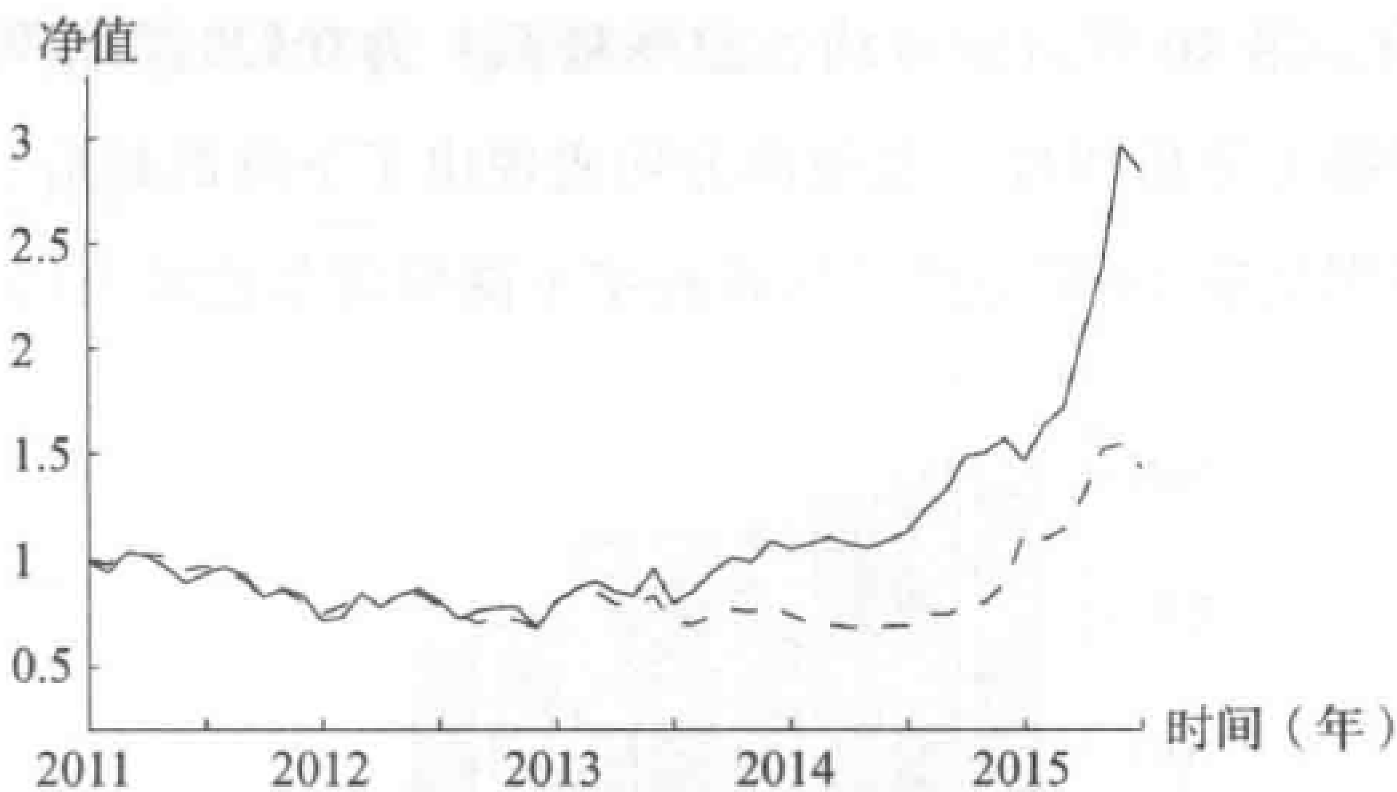


图 5-16　反转因子选股前 1/3 的股票组合净值走势

图 5-17 是逐月地将反转因子排序前 1/3 的股票组合收益率减去沪深 300 指数的收益率后，所得到的收益率序列形成的净值走势情况。净值从 2010 年 12 月收盘的 1 开始，在 2015 年 6 月底达到 1.8439，四年半的收益率为 84.39%，折算下来的年化收益率为 14.57%。四年半之间的最大回撤率为 38.43%，回撤的开始点为 2014 年 9 月收盘的 1.9495，结束点则是 2014 年 12 月收盘的 1.2003，一共三个月的回撤期与市值因子 3 等分的情况完全一致。

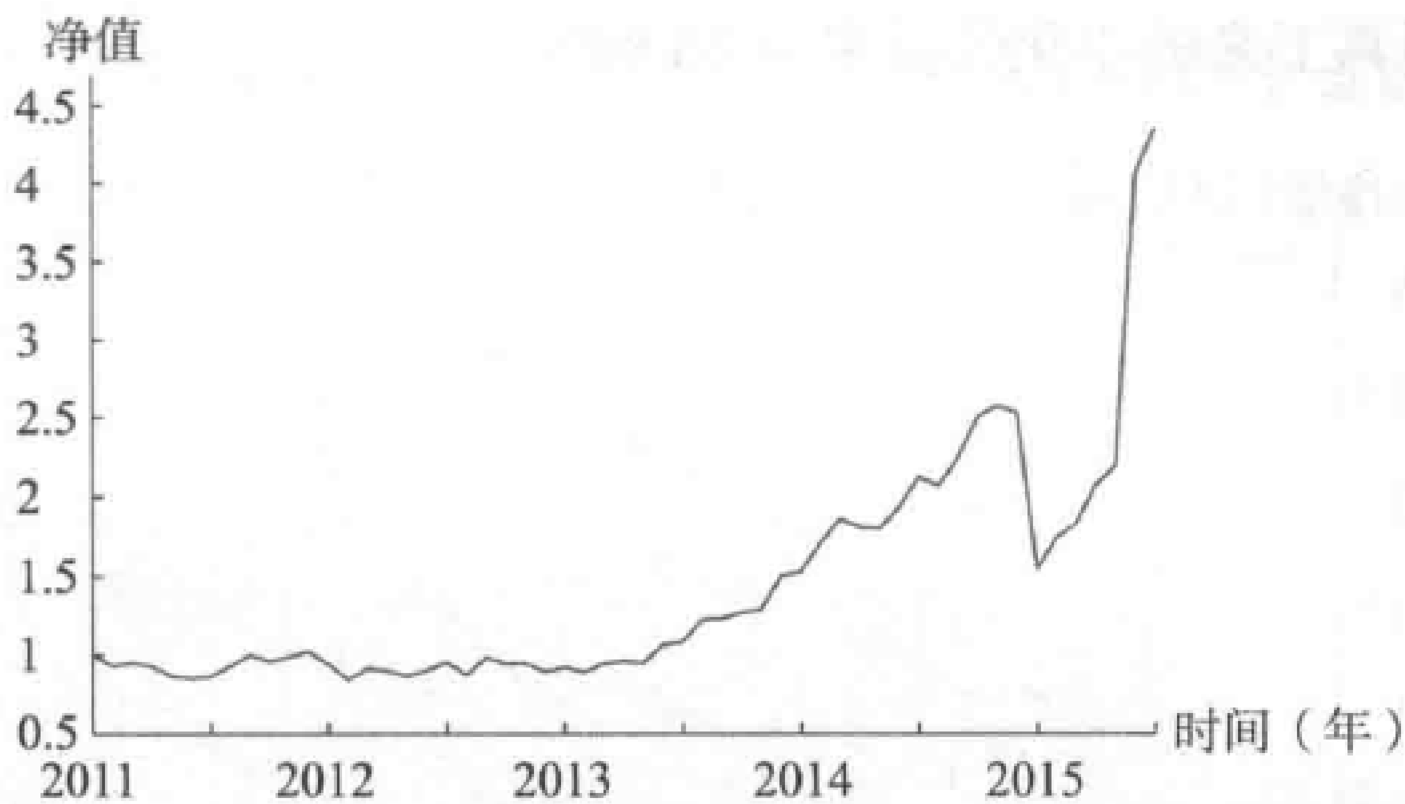


图 5-17　反转因子选股前 1/3 的股票组合与沪深 300 收益之差的净值走势

5.3.2　反转因子选股 10 等分

图 5-18 展示了将所有股票按照上月收益率大小划分为 10 等份，每一等份中的股票又按照等权重进行组合时，10 个等份各自的平均月度收益情况。第 1 等份的月均收益率为 2.00%，而第 2 等份的月均收益率略高，为 2.01%，也是 10

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

个等份中最高的，第 10 等份的月均收益率最低，为 0.12%。整体上来看，除了第 2 等份略高于第 1 等份以外，其他部分均表现出了下降的趋势。相对而言，最后两个等份，特别是第 10 等份的月均收益率下降幅度是比较大的。

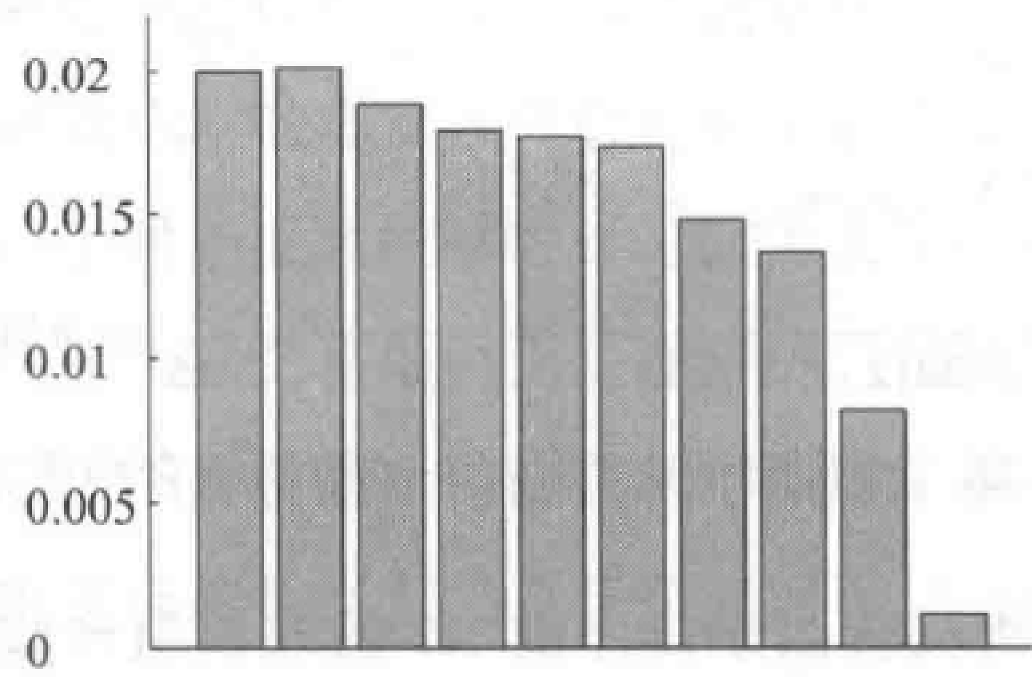


图 5-18　反转因子选股 10 等分的收益情况

在确定了逐月的排序和 10 等分之后，也就得到了通过上个月收益反转因子选出的第 1 等份股票组合的逐月变动详情。图 5-19 中的实线展示了使用反转因子选股策略进行交易时，以 1 为初始值的净值走势情况，虚线则是用来对比的沪深 300 指数走势。策略净值在 2015 年 6 月底达到 2.9113，四年半的策略收益率为 191.13%，折算下来的年化收益率为 26.80%。

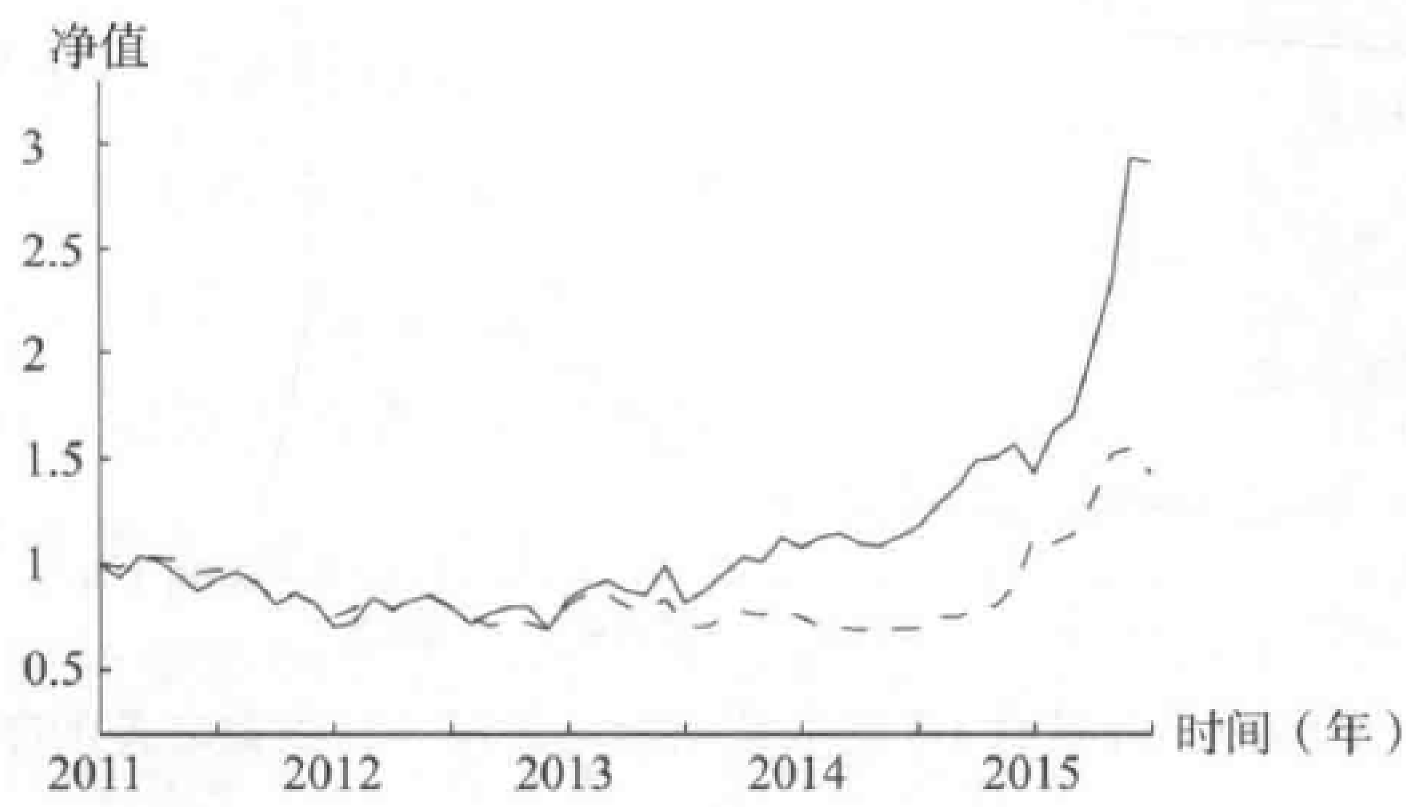


图 5-19　反转因子选股前 1/10 的股票组合净值走势

图 5-20 则是逐月的将反转因子排序前 1/10 的股票组合收益率减去沪深 300 指数的收益率后，所得到的收益率序列形成的净值走势。净值从 2010 年 12 月收盘的 1 开始，在 2015 年 6 月底达到 1.8792，四年半的收益率为 87.92%，折算下来的年化收益率为 15.05%。四年半之间的最大回撤率为 40.46%，回撤的开始点

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

为 2014 年 9 月收盘的 1.9684，结束点则是 2014 年 12 月收盘的 1.1720。

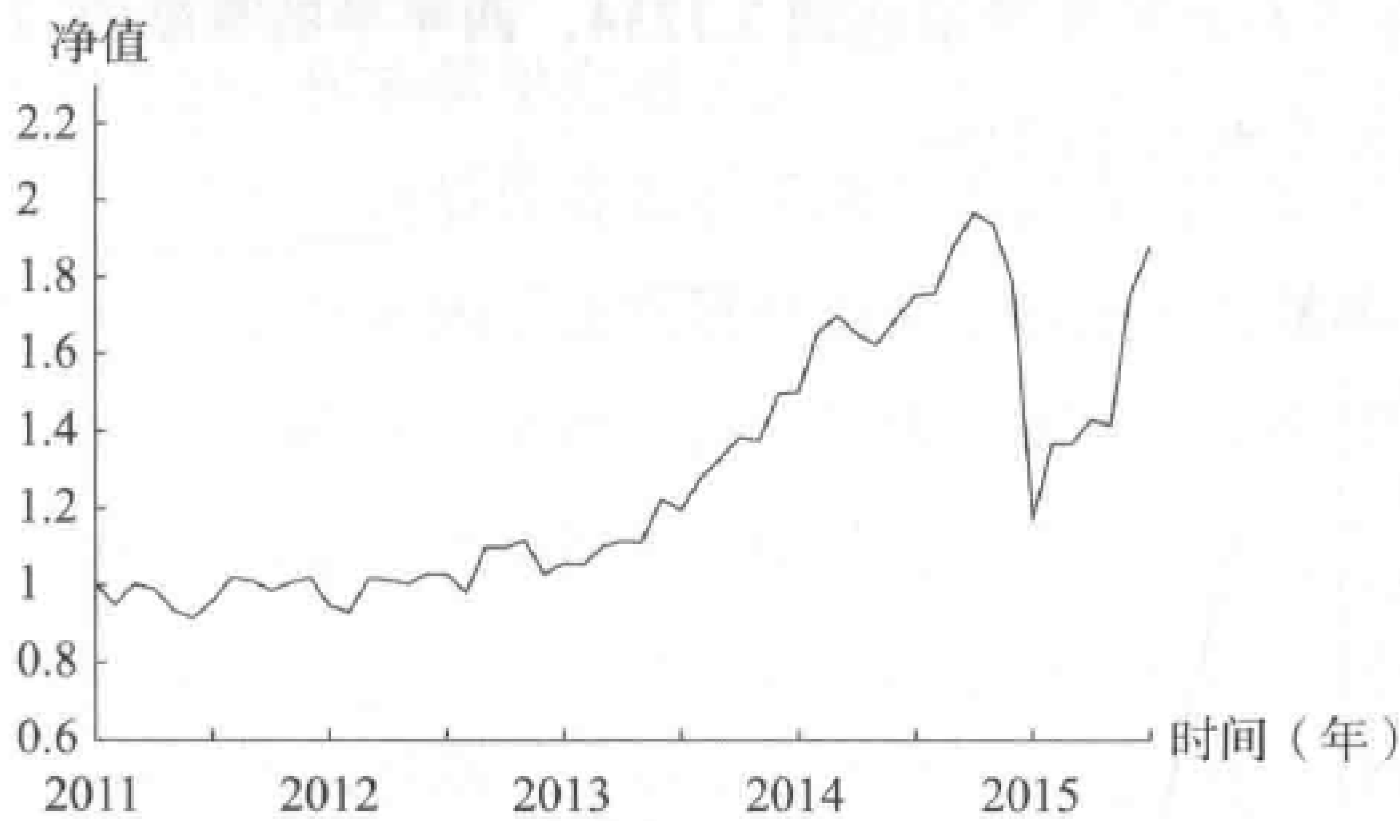


图 5-20　反转因子选股前 1/10 的股票组合与沪深 300 收益之差的净值走势

5.3.3　反转因子选股 30 等分

图 5-21 展示了将所有股票按照上个月收益率大小划分为 30 等份，每一等份中的股票又按照等权重进行组合时，30 个等份各自的平均月度收益情况。第 1 等份的月均收益率最高，为 2.19%，第 30 等份最低，为 −0.75%。第 29 等份和第 30 等份之间的月均收益率相差 1.32%，相对较大，在图中也能看到一个非常明显的收益率由正转负的变化。与此同时，最后几个等份展现了比前面等份更为明显的下降趋势。

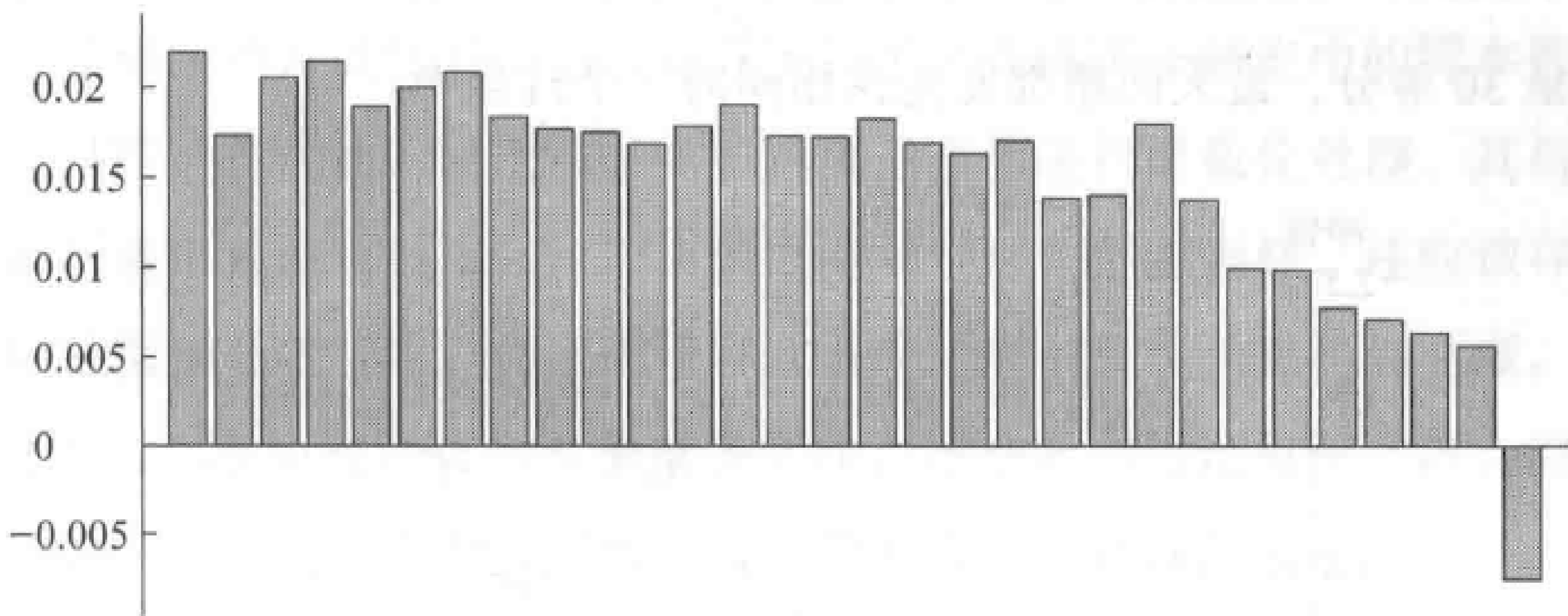


图 5-21　反转因子选股 30 等分的收益情况

在确定了逐月的排序和 30 等分之后，也就得到了通过反转因子选出的第 1 等份股票组合的逐月变动详情。图 5-22 中的实线展示了 30 等分的情况下，以 1

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

为初始值的反转因子选股策略净值走势，虚线则是用来进行比较的沪深 300 指数走势。到 2015 年 6 月底策略净值达到 3.2234，四年半的策略收益率为 222.34%，折算下来的年化收益率为 29.71%。

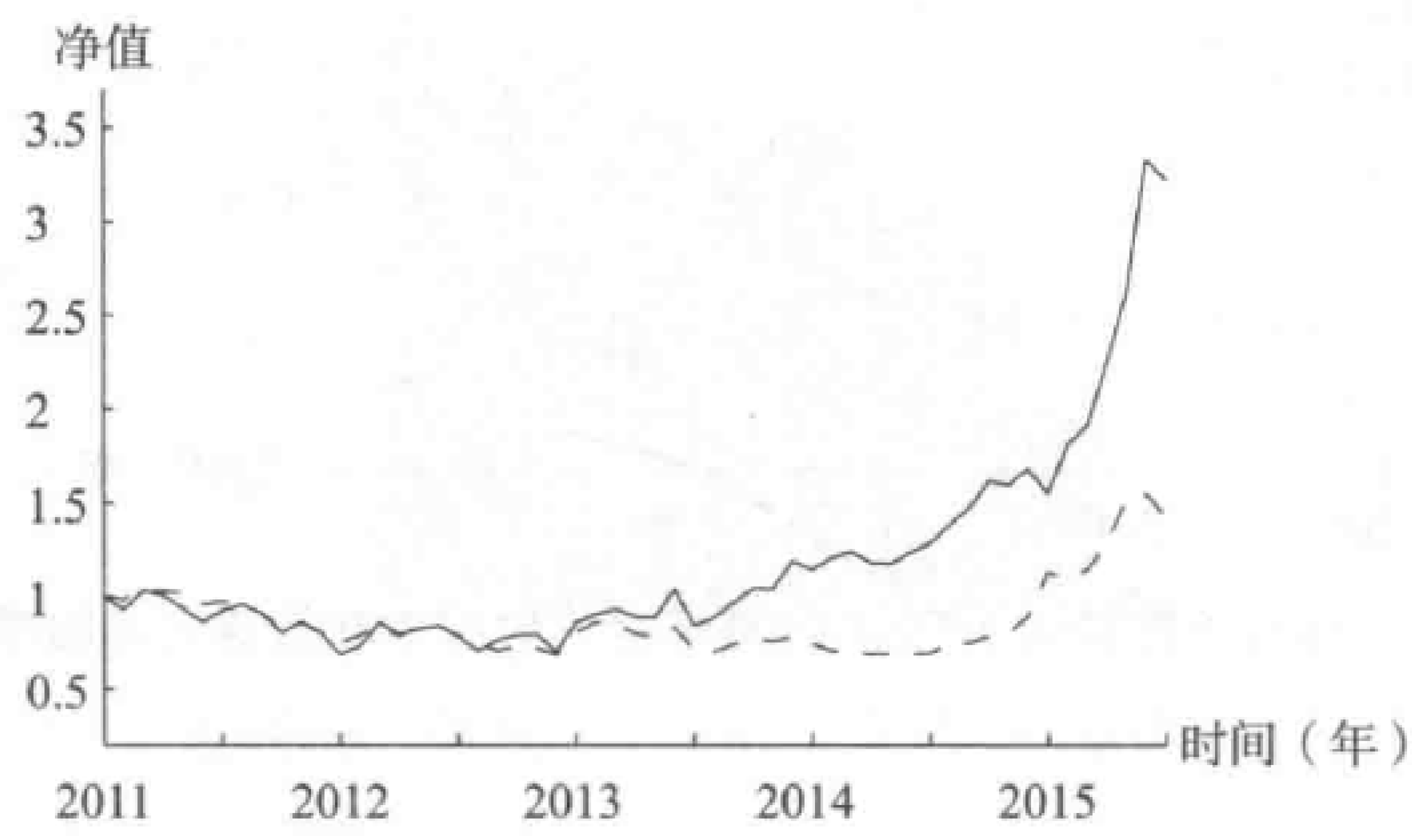


图 5-22 反转因子选股前 1/30 的股票组合净值走势

图 5-23 则是逐月的将上个月收益反转因子排序前 1/30 的股票组合收益率减去沪深 300 指数的收益率后，所得到的收益率序列形成的净值走势情况。净值从 2010 年 12 月收盘的 1 开始，在 2015 年 6 月底达到 2.1051，四年半的收益率为 110.51%，折算下来的年化收益率为 17.99%。四年半之间的最大回撤率为 40.19%，回撤的开始点为 2014 年 9 月收盘的 2.1441，结束点则是 2014 年 12 月收盘的 1.2824。可以发现，不同于市值因子的情况，反转因子不论是 3 等分、10 等分还是 30 等分，最大回撤都发生在相同的三个月当中。

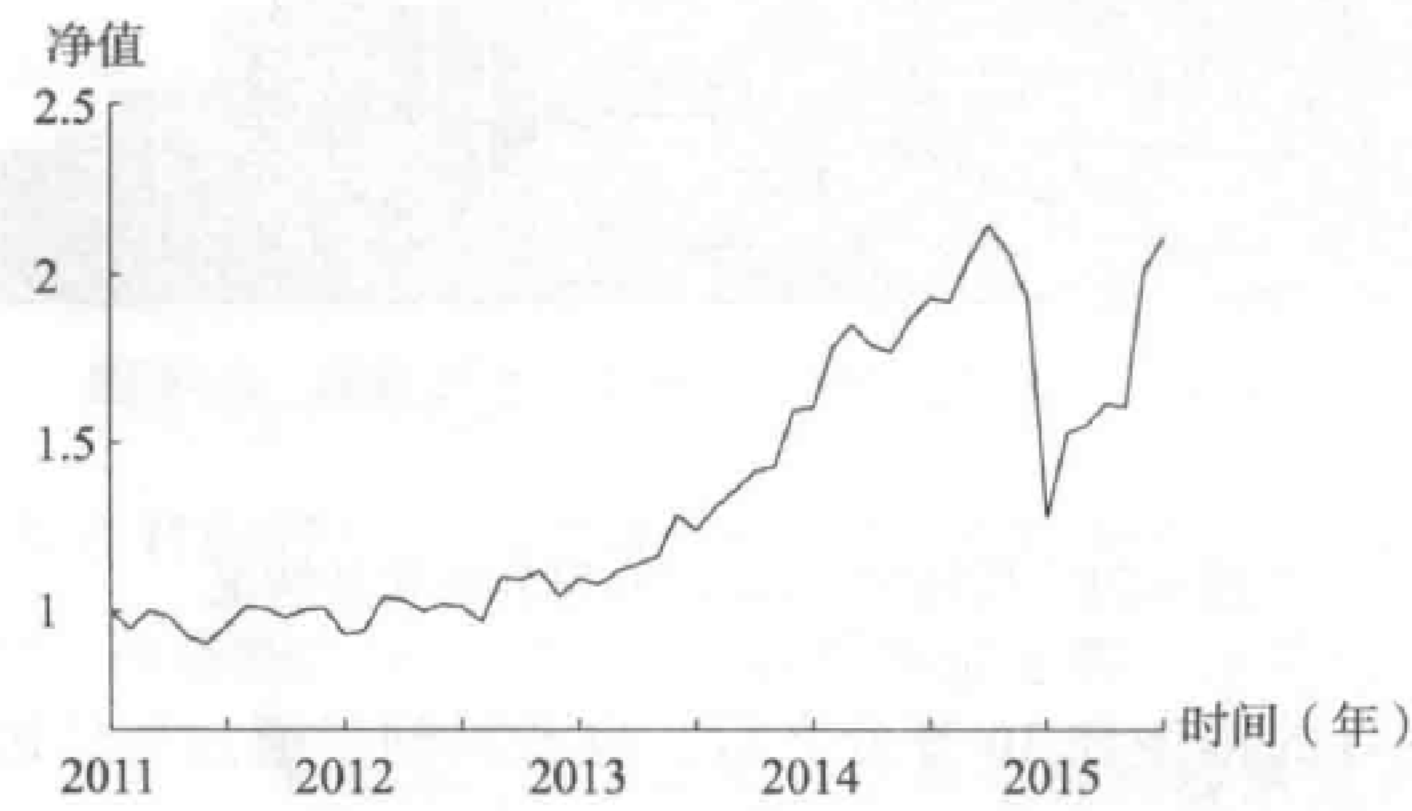


图 5-23 反转因子选股前 1/30 的股票组合与沪深 300 收益之差的净值走势

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

5.4 多因子选股策略的简单优化

在本章开头的部分，作者简单介绍了多因子选股策略是如何在若干个有效因子的基础之上进行构建的。随后作者展示了两个因子的实际案例，可以看到，流通市值因子和上个月收益反转因子在所研究的样本数据之上是两个具有预测效力和盈利能力的选股因子。在这一节中，作者就将这两个因子组合成多因子模型并用来选择股票，这样的做法一来遵循循序渐进的过程，二来也简化了多因子策略的复杂度，用最基本的两个选股因子来做直观的案例展示，从而方便读者更好地了解多因子选股策略的整体框架。需要注意的是，本书中使用的两个选股因子都在 2014 年年底和 2015 年年初出现了比较极端的失效情况，因此多因子选股策略会相应地经历一个短暂但程度较大的回撤也是必然的结果，这在本节和后面章节中的实际案例里都会有具体的体现。不过就多因子选股策略的本质而言，有效因子越多、因子配置越均衡，多因子选股策略受到单因子失效影响的程度也就会越小，这其实和投资组合配置的情况是类似的，实际交易中研究人员还是应该尽可能地多发掘出一些相关性较小的选股因子，从而增强多因子选股策略的稳健性。

为了研究多因子选股策略，作者获取的同样是自 1991 年 1 月开始直到 2015 年 6 月底结束的国内股票月度数据。所选股票的范围涵盖在深圳证券交易所进行交易的所有 A 股，包括主板、中小企业板和创业板三个板块，以及在上海证券交易所进行交易的所有 A 股。月度数据则包括月末收盘时股票的复权后月度收益率，以及股票的月度流通市值，这其实与单个市值因子研究中的样本数据完全一致。但是在进行线性回归之前，需要对流通市值进行对数化处理，其目的是为了更好地反映出月度收益率和流通市值之间的线性关系。当然，还应该存在着更为适合的数据预处理手段，使得因子和收益率之间的线性关系更加明晰。不过这属于第 2 章 2.2 节中提到过的数据准备工作，不在本书讨论范围之内，因此不再做深入研究，读者可以在研究过程中自行摸索合理的处理方法。

开始挑选股票进行买卖的时间同样从 2011 年 12 月收盘开始，也就是说在回溯测试样本中将一共交易四年半的时间。1991 年 1 月到 2010 年 12 月的所有月度流通市值数据和月度收益率则作为开始买卖股票前的备用数据，不过对于本节中简单的多因子选股策略而言也只需要用到 2010 年 12 月这一个月的数据。而

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

在整个交易过程中，回溯动作也设定为只进行买入和卖出操作，不进行卖空的操作，资金全部用于持有股票，与单因子的研究环境一致。

回溯测试的过程如下，在第一个选股和交易的时间点，即2010年12月收盘时，将此时所有股票的月收益率、流通市值与2011年1月的收益率进行匹配，并进行回归估计，如下所示：

$$r_{2011-1}=a+b_1S_{2010-12}+b_2r_{2010-12}$$

式中，r_t是2011年1月的月收益率；S_{t-1}是2010年12月收盘时股票的流通市值；r_{t-1}是2010年12月的月收益率，通过估计该回归式可以得到针对2011年1月的a、b_1、b_2三个参数估计值。

首先将时间按月份向后推进，每一个月都可以得到一组参数估计值，直到2015年6月收盘时结束，一共54组估计参数；然后将54个a进行平均得到$\overline{a}$，将54个b_1平均得到$\overline{b}_1$，将54个b_2平均得到$\overline{b}_2$，那么基于整个样本所得到的最终回归结果就是

$$r_t=\overline{a}+\overline{b}_1S_{t-1}+\overline{b}_2r_{t-1}$$

同第4章中所介绍的简单择时策略一样，作者在估计完回归式之后，又会将该回归式用到相同的样本数据上进行交易的模拟。虽然这样会产生未来信息的问题，但是相对简单的流程处理可以方便读者进行理解。具体来说，就是在2010年12月收盘的选股时间点上，基于2010年12月所有股票的月度流通市值和月度收益率数据，使用最终得到的回归式计算得到所有股票在2011年1月的收益率的预测值。其次将股票按照收益率预测值从高到低进行排序，再按顺序等分为若干个部分，每一等份中的股票又按照等权重进行组合，也就是买入占用的资金量相等的组合。然后在2011年2月底进行相同的排序和等分过程，各个等份中的股票组成基于预测收益率排序的变化而发生改变。之后每一个月收盘时都进行相同的操作，直到交易样本结束即2015年6月收盘为止。

排序完成之后的处理，与单因子的研究内容完全一致。首先根据每一个等份下长度为54的收益率序列，计算出各个等份的平均月度收益情况，进而判断多因子模型预测下一期收益率的整体趋势性。不过由于多因子模型本身的作用不在于挑选合适的因子，因此这一部分的结果仅作参考，实际意义小于其在单因子

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

研究中的作用。交易策略方面，则根据排序中的第一个等份选择股票，这也就是每一个月多因子选股策略所选择的股票组合。根据这个长度为 54 的收益率序列，就可以计算出多因子选股策略的净值走势。需要指出的是，多因子选股模型预测的是下一个月各只股票的收益率水平，因此按预期收益从大到小排序，排序第一的股票就应该是建仓的选择，这是符合交易逻辑的。而本书中挑选出来的两个单因子恰好都是收益预期与因子大小具有反向关系，因此按照因子从小到大进行排序，建仓选择才是排序第一的股票组合。希望读者能够区分这之间的差别。

和单因子的研究设置相仿，作者将所有股票按照预期收益大小排序划分为 3 等份、10 等份、30 等份共三种情况来加以研究。

5.4.1 多因子选股 3 等分

图 5-24 展示了将所有股票按照多因子模型的预期收益大小划分为 3 等份，每一等份中的股票又按照等权重进行组合时，3 个等份各自的平均月度收益情况。具体的收益率大小分别为 2.43%、1.53% 和 0.53%。等份之间的差额分别为 0.9% 和 1%，呈现出一个比较稳定的下降趋势。

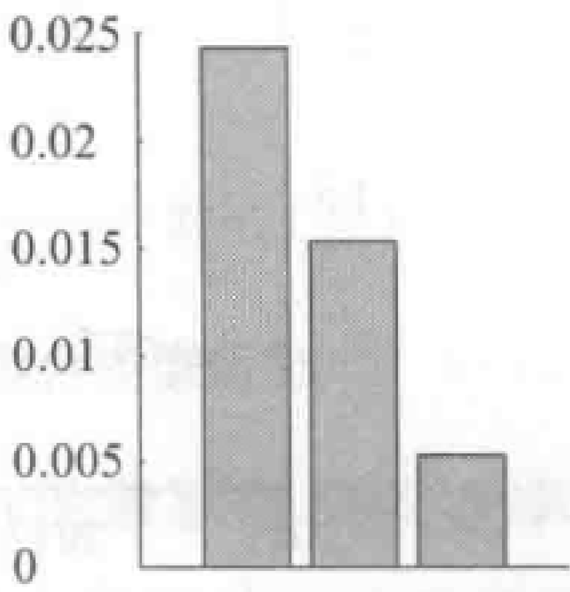


图 5-24 多因子选股 3 等分的收益情况

确定了逐月的排序和 3 等分之后，就可以得到第 1 等份股票组合的逐月变动情况，也就是多因子选股策略的交易详情。图 5-25 中的实线展示了使用多因子选股策略进行交易的情况下，以 1 为初始值的策略净值走势，虚线则是用来进行对比的同期沪深 300 指数走势。策略净值在 2015 年 6 月收盘时达到 3.6539，四年半的策略收益率为 265.39%，折算下来的年化收益率为 33.37%。

图 5-26 则是逐月地将多因子选股排序前 1/3 的股票组合收益率减去沪深 300 指数的收益率后，所得到的收益率序列形成的净值走势。净值从 2010 年 12 月收盘的 1 开始，在 2015 年 6 月底达到 2.3364，四年半的收益率为 133.64%，折算下来的年化收益率为 20.75%。四年半之间的最大回撤率为 40.50%，回撤的开始点为 2014 年 9 月收盘的 2.3757，结束点则是 2014 年 12 月收盘的 1.4136，一共

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

三个月的回撤期与两个单因子的研究一致。

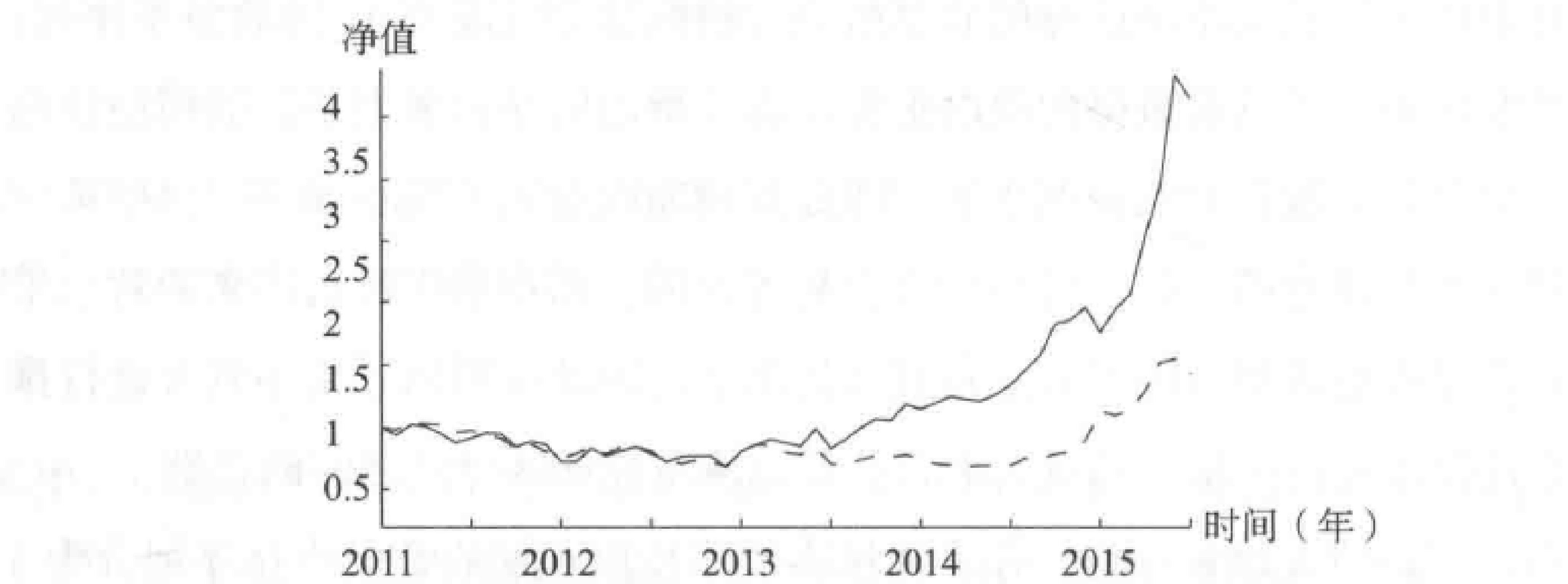


图 5-25　多因子选股前 1/3 的股票组合净值走势

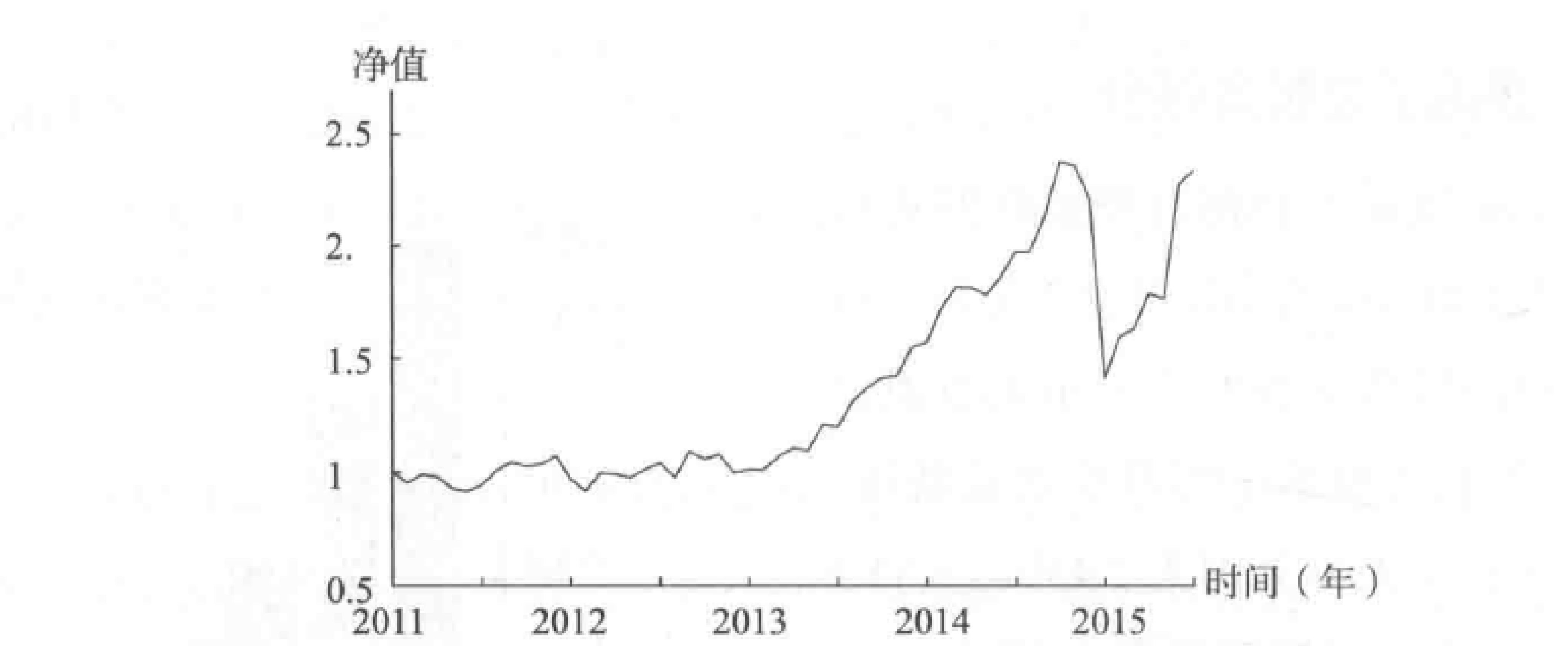


图 5-26　多因子选股前 1/3 的股票组合与沪深 300 收益之差的净值走势

5.4.2　多因子选股 10 等分

图 5-27 展示了将所有股票按照多因子模型的预期收益大小划分为 10 等份，每一等份中的股票又按照等权重进行组合时，10 个等份各自的平均月度收益情况。第 1 等份的月均收益率是 10 个等份中的最高值，为 2.78%，第 10 等份的月均收益率则是 10 个等份中的最低值，为 0.14%。其他等份也按照排序关系存在严格的下降趋势，而且下降的幅度也比较

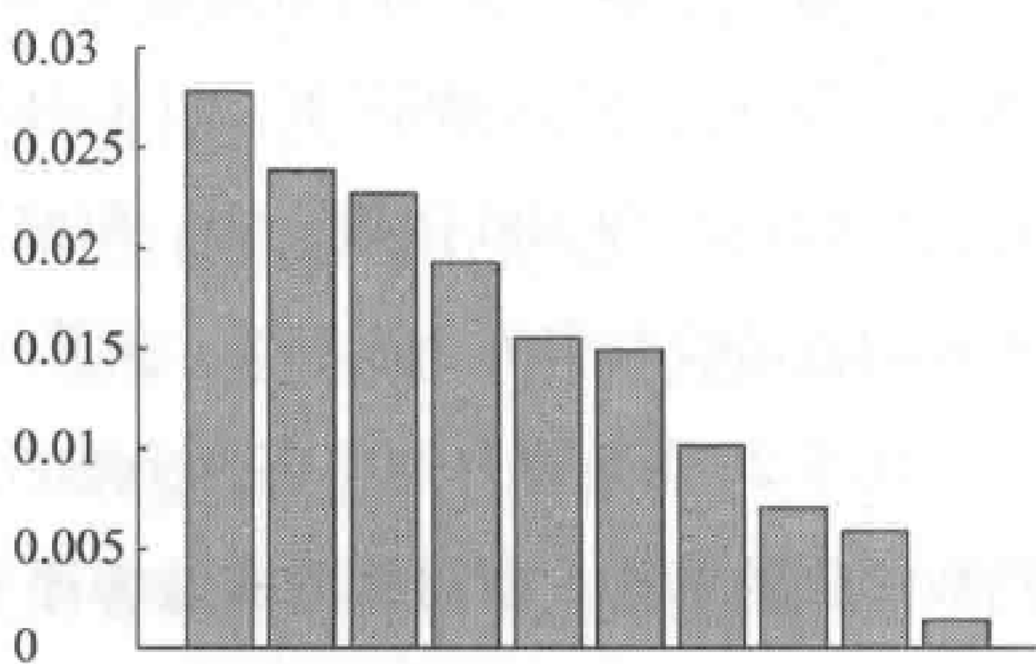


图 5-27　多因子选股 10 等分的收益情况

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

稳定。与前两节中的单因子研究对比起来，可以发现多因子模型使得等分下的未来收益率具有更平稳的趋势。

在确定了逐月的排序和 10 等分之后，也就得到了通过多因子模型选出的第 1 等份股票组合的逐月变动详情。图 5-28 中的实线展示了使用多因子选股策略进行交易时，以 1 为初始值的净值走势情况，虚线则是用来对比的沪深 300 指数走势。策略净值在 2015 年 6 月底达到 4.3883，四年半的策略收益率为 338.83%，折算下来的年化收益率为 38.91%。

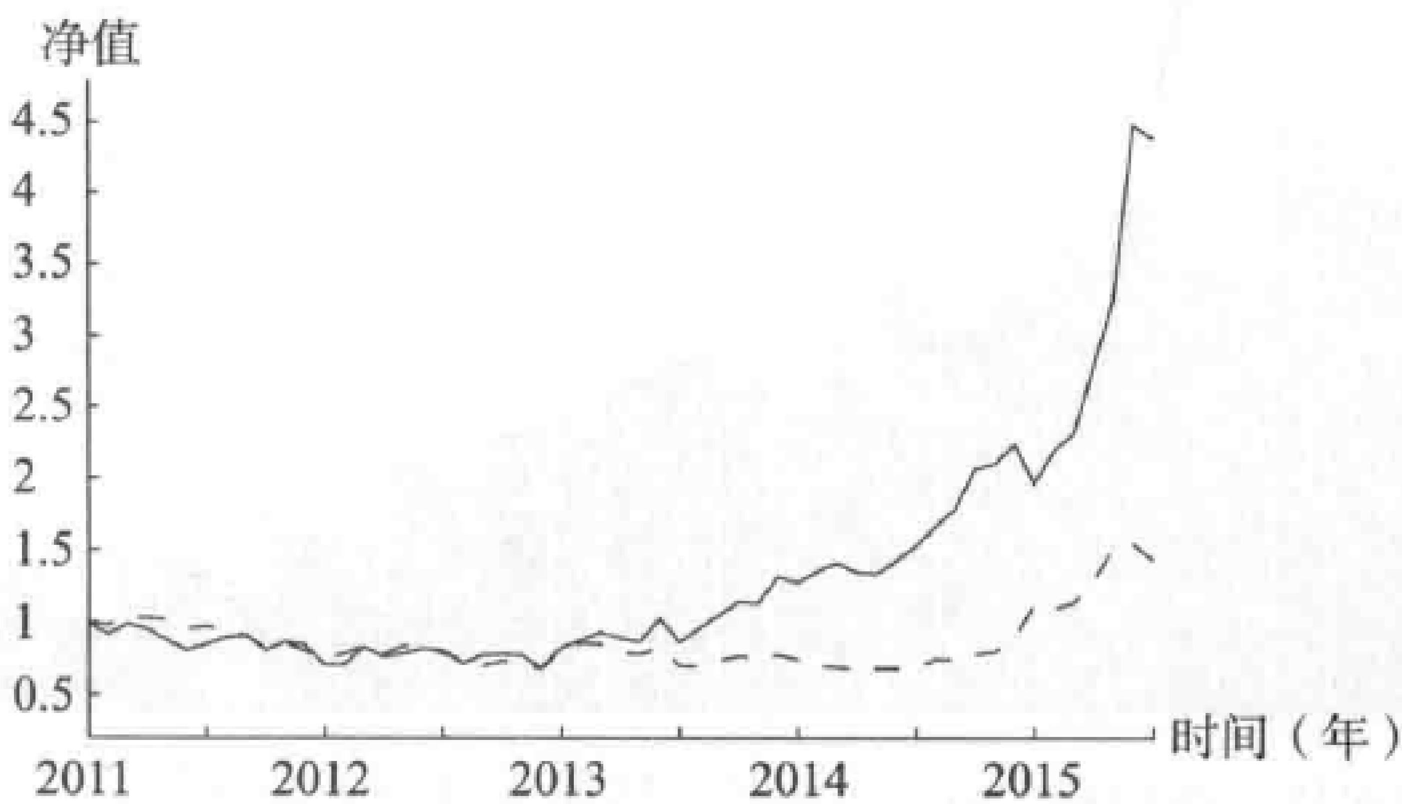


图 5-28 多因子选股前 1/10 的股票组合净值走势

图 5-29 则是逐月地将多因子选股排序前 1/10 的股票组合收益率减去沪深 300 指数的收益率后，所得到的收益率序列形成的净值走势。净值从 2010 年 12 月收盘的 1 开始，在 2015 年 6 月底达到 2.7747，四年半的收益率为 177.47%，折算下来的年化收益率为 25.46%。四年半之间的最大回撤率为 41.79%，回撤的开始点为 2014 年 9 月收盘的 2.6863，结束点则是 2014 年 12 月收盘的 1.5638。

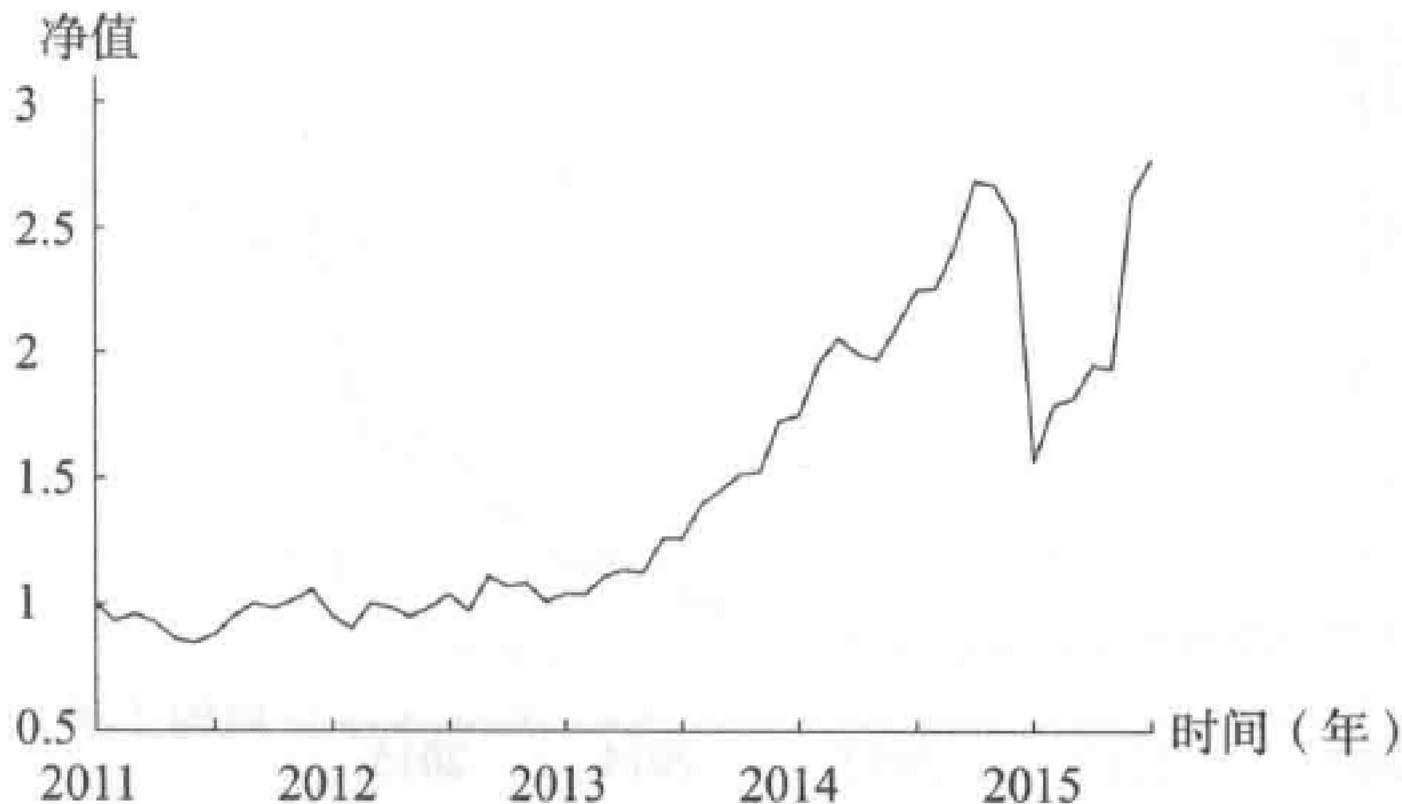


图 5-29 多因子选股前 1/10 的股票组合与沪深 300 收益之差的净值走势

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

5.4.3 多因子选股 30 等分

图 5-30 展示了将所有股票按照多因子模型的预期收益大小划分为 30 等份，每一等份中的股票又按照等权重进行组合时，30 个等份各自的平均月度收益情况。第 1 等份的月均收益率同样也是 30 个等份中的最高值，为 2.89%；第 30 等份的月均收益率也是 30 个等份中的最低值，为 −0.25%。其他等份的收益情况也基本显示出一个比较平稳的下降趋势，同样对比前两节中的单因子研究可以发现，多因子模型使得等分下的未来收益率具有更平稳的趋势。

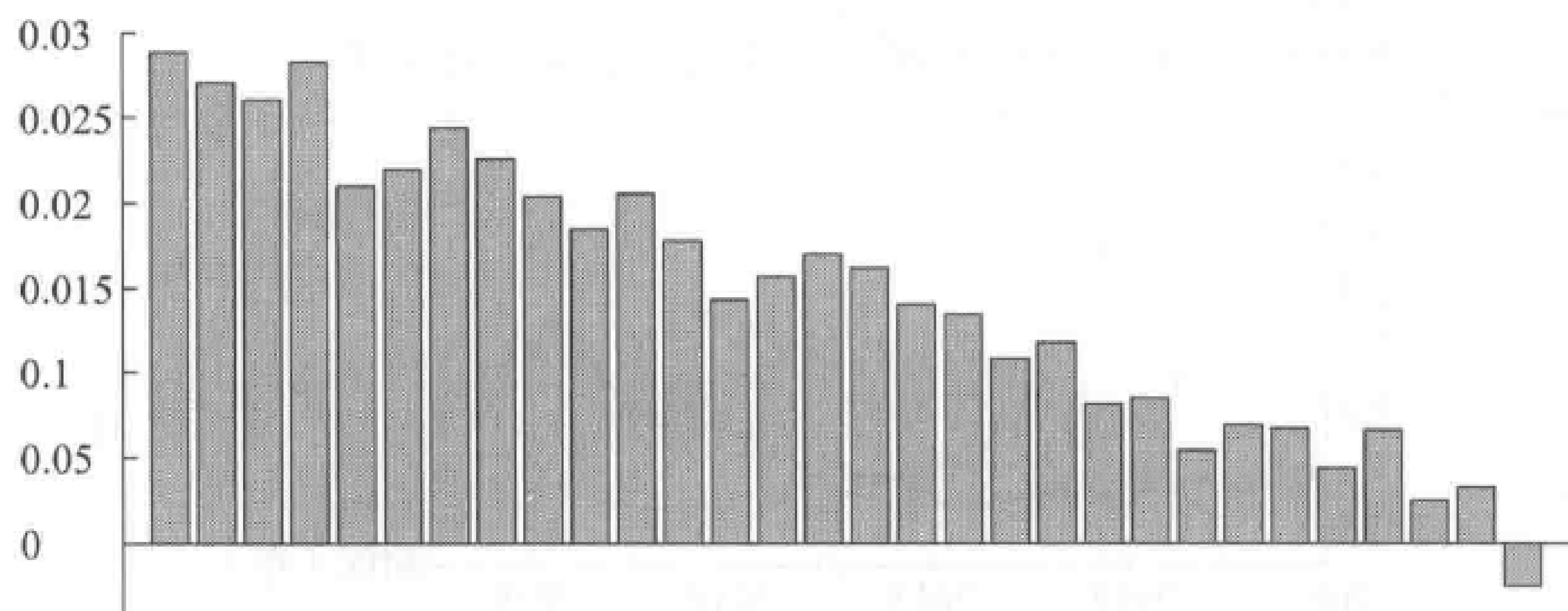


图 5-30 多因子选股 30 等分的收益情况

在确定了逐月的排序和 30 等分之后，也就得到了通过多因子模型选出的第 1 等份股票组合的逐月变动详情。图 5-31 中的实线展示了 30 等分的情况下，以 1 为初始值的多因子选股策略净值走势，虚线则是用来进行比较的沪深 300 指数走势。策略净值在 2015 年 6 月底达到 4.6520，四年半的策略收益率为 365.20%，折算下来的年化收益率为 40.72%。

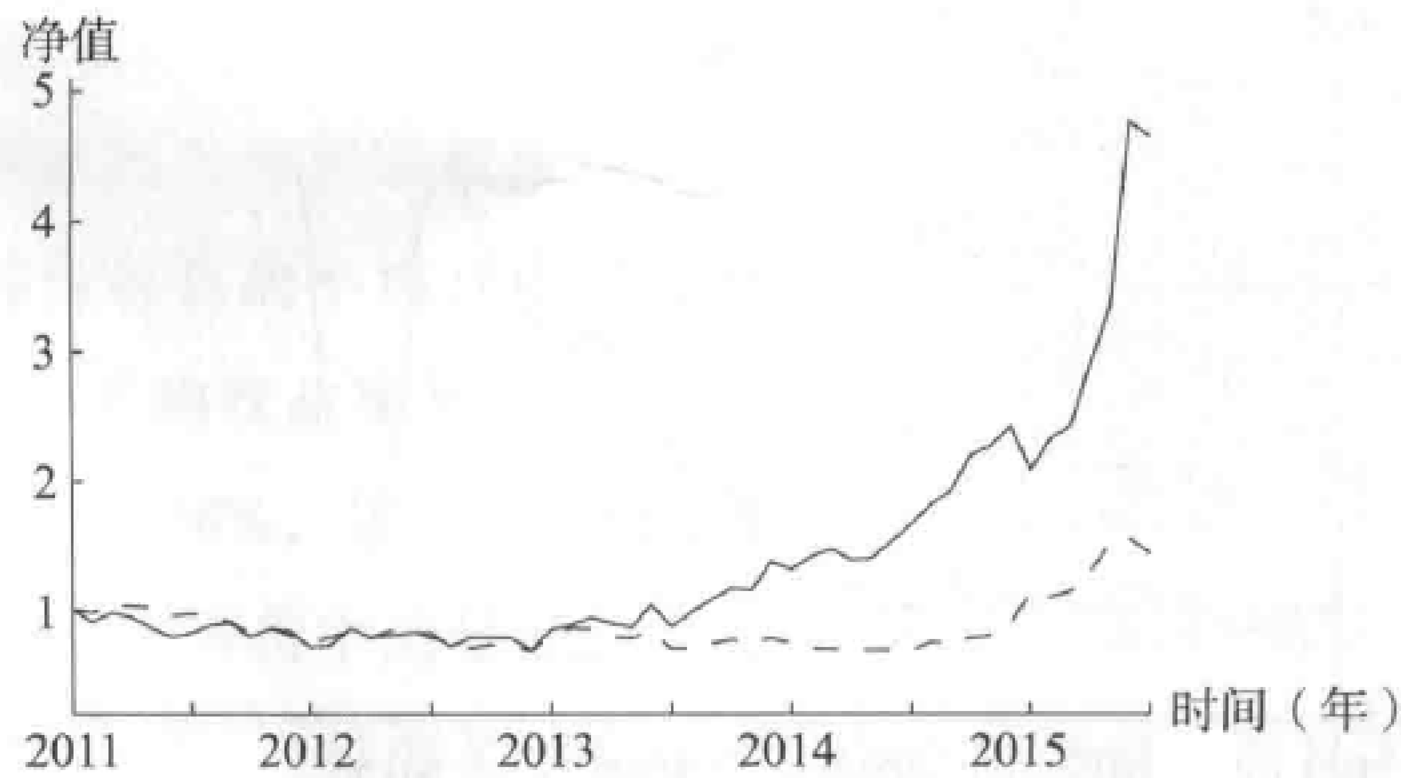


图 5-31 多因子选股前 1/30 的股票组合净值走势

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

图 5-32 则是逐月地将多因子选股排序前 1/30 的股票组合收益率减去沪深 300 指数的收益率后，所得到的收益率序列形成的净值走势。净值从 2010 年 12 月收盘的 1 开始，在 2015 年 6 月底达到 2.9179，四年半的收益率为 191.79%，折算下来的年化收益率为 26.87%。四年半之间的最大回撤率为 43.08%，回撤的开始点为 2014 年 10 月收盘的 2.9003，结束点则是 2014 年 12 月收盘的 1.6508，回撤期的长度为两个月。

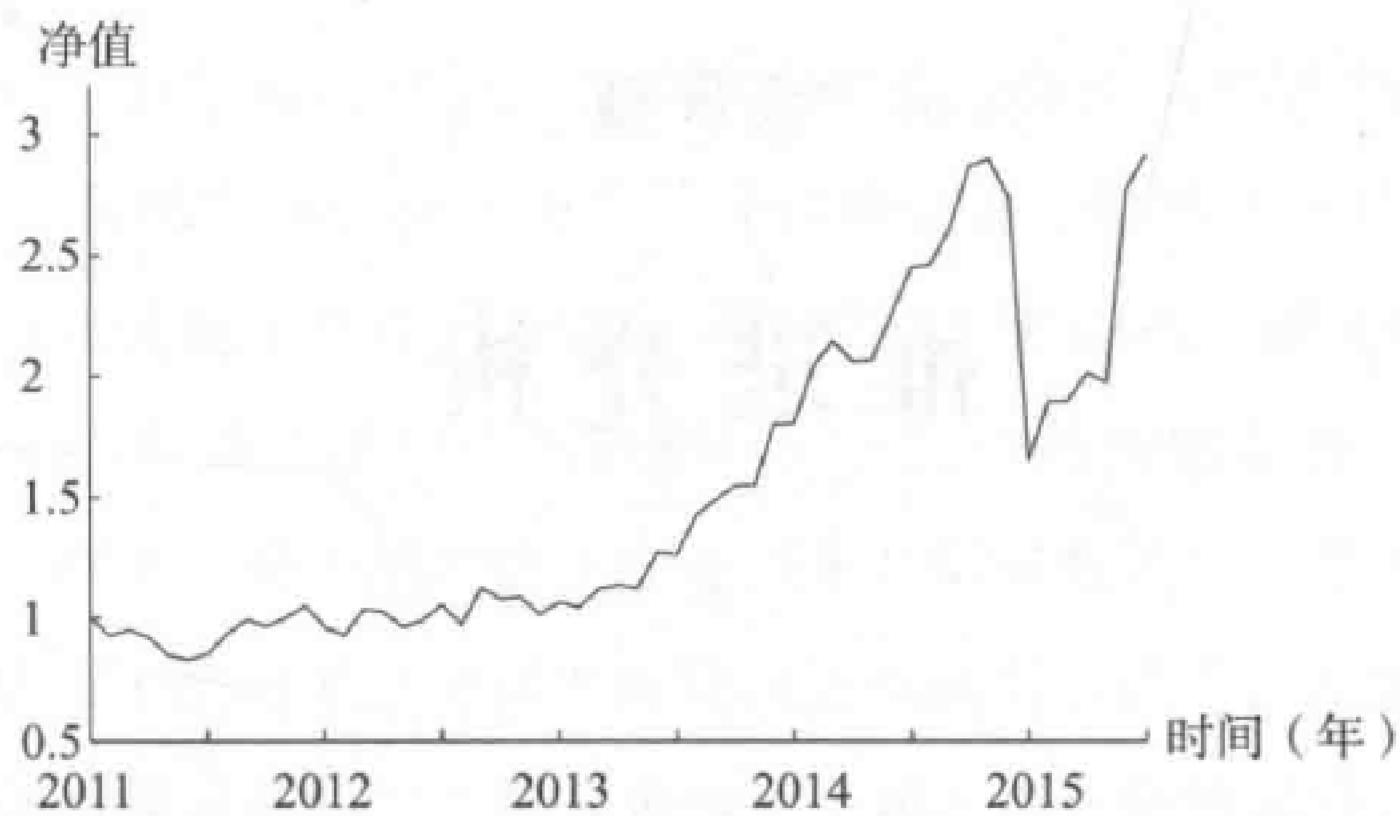


图 5-32　多因子选股前 1/30 的股票组合与沪深 300 收益之差的净值走势

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

第6章

推进分析

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

6.1 推进分析框架

在第 4 章和第 5 章的策略研究当中，由于回溯测试过程简单化的处理，因此不可避免地会受到未来信息问题的影响。与此同时，最优化的处理只在整个样本数据上进行了一次，因此如果只分析回溯测试的结果，研究人员无法判断研发得到的具体策略是否具有过度拟合的问题。

在这一章中，作者就将介绍一种具有实践意义的分析方法来处理这些问题，也就是章节名称所提到的推进分析。在具体说明推进分析的框架和使用方法之前，本节首先介绍一下与推进分析有着极高关联性但是适用范围更加广泛的样本外检验，方便读者循序渐进地了解推进分析的框架结构，更好地掌握这种研究方法。

样本外检验是一种被广泛应用于数据科学研究中的检验方法。在数据科学的一般研究情况下，样本外检验的功能主要是用来判断对数据的处理方法是否存在明显的过度拟合。图 6-1 给出了一个样本外检验的示例用以说明，先忽略下面的时间轴，实际上样本外检验就是将整个样本分为两个相互不重叠的部分，也就是白色的样本内和灰色的样本外。首先针对样本内的数据进行模型的研究和优化，得到一个在样本内数据上最优的模型方法，再将已经优化完成的模型方法直接放入样本外的数据进行检验，如果样本外的数据检验结果依然说明该模型方法具有显著的效果，那么就可以在一定程度上判断出该模型方法对实际问题处理的有效性。对于在样本内优化得到的具体模型方法，比较其在样本内和样本外数据上各自的效果，两者越接近，那么该模型方法在样本内进行优化时的过度拟合问题就越小。

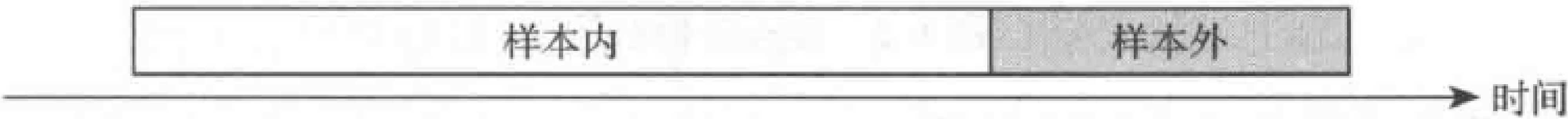


图 6-1 样本外检验示例

在样本内、样本外的示例下面加上一个时间轴，是为了指出在处理资产价格等具有时间特征的数据时，往往将样本内的部分取在时间段前期，而将样本外的部分取在时间段后期。当然，将样本外的部分放在时间轴的前面，甚至将样本外分为两部分置于样本内的前后两端，都是可以接受的处理方法，符合样本外检验的基本定义。但是相比较而言，图 6-1 的样本处理方式更加合乎逻辑，更为贴合

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

时间流逝下的数据特征。因此一般而言，这是一种更被接受的处理方法。

量化交易策略的研发作为数据科学的一个专业化分支，实际上是可以直接使用样本外检验来进行过度拟合判断的。但是正是由于量化交易策略研发作为一个分支，具有自身的特殊性和专业性，因此存在一些更为合适的研究方法来判断过度拟合的问题。推进分析正是其中比较具有代表性的一种方法，其不仅可以判别量化交易策略的过度拟合程度，同时具有其他一些比较好的性质特点，是量化交易策略研究中比较常见且值得使用的检验方法。

图 6-2 给出了一个推进分析的示例。最下面的数字代表随时间前进的交易周期，例如在第 4 章量化择时策略的数据下，每个数字都代表 1 日，在第 5 章量化选股策略的数据下，每个数字都代表 1 个月，而如果使用的样本数据为分笔数据的话，那么每一个数字都代表两次分笔数据到达的间隔。也就是说，推进分析的计算周期或者说推进周期应该以交易判断的间隔为准，而不是以某一个固定的时间长度为准，日度数据和月度数据的例子只是因为判断交易的时间间隔恰好是日和月而已。

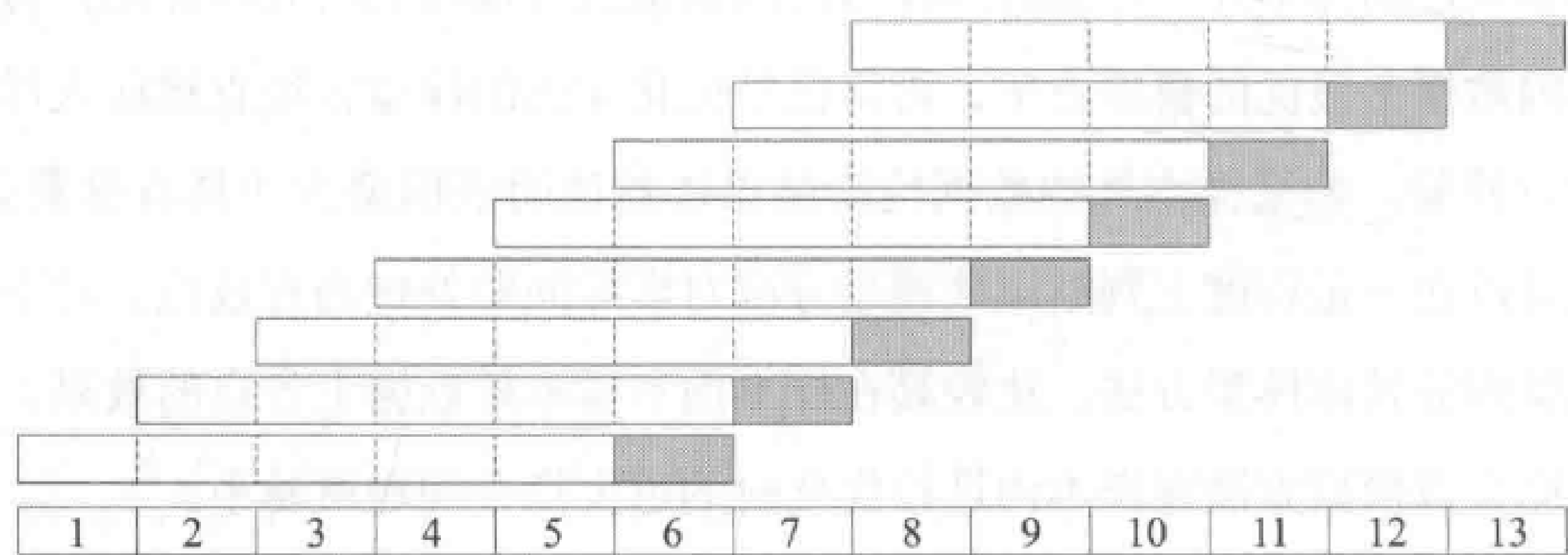


图 6-2　推进分析示例

首先来看图 6-2 中离计数条最近的一行白色框加灰色框，这里的操作类似图 6-1 中的样本外检验过程。其中白色框包括从周期 1 到周期 5 的数据，灰色框则是周期 6 的数据。假设研究人员要检验的是一个参数待估计的量化交易策略，现在先基于白色框的数据对其进行最优化，得出相应的参数，然后将参数确定好的策略放入灰色框的数据中进行检验，得到相应的结果。优化的具体目标既可以是量化交易策略的整体收益情况，也可以是一个具体模型的预测能力，这在前面的说明和案例讲解中均有涉及，这里不再多做赘述。可以发现，针对这一行数据的

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

研究操作实际上就是一个样本外检验的过程，但是该检验具有一个明显的特点，就是样本外的数据长度仅为一个交易周期。

随后的操作基本上就是推进分析这个名称的由来。将样本内和样本外两个部分分别按照时间前进的方向向后推进一个周期，白色框的样本内数据变为了周期2到周期6的数据，灰色框的样本外数据则变成了周期7的数据。同样的，先基于白色框的数据对其进行最优化，得出相应的参数，然后将参数确定好的策略放入灰色框的数据中进行检验，得到对应的结果，这也就是第二行数据得到的样本外检验结果。

之后则每次推进一个周期，进行相同的检验操作，如图6-2中的上面6排，每一行数据都有一个周期下的样本外检验结果。最后将所有灰色框的检验结果汇总，就可以得到从第6个周期直到结束的推进分析检验结果，作者也称其为全部样本下的推进分析检验结果。

值得注意的是，如果具体的量化交易策略本身需要备用数据来参与计算，那么整个研究实际需要的数据量将超过计数框的限制，提前到周期1之前。灰色框的检验计算也需要用到白色框中的数据，从而完成策略的计算。这些内容在图6-2中都没有显示出来，希望读者自行理解。

从上面对推进分析的介绍中其实不难发现，推进分析的整个操作流程非常贴合实际交易的情况。在实际环境中如果使用量化交易策略进行交易，使用者常常会先根据最近的数据优化好策略设置和策略参数之后，再用其来指导当前的交易，而在下一次需要判断是否交易时，再重新优化一遍策略设置和策略参数，从而使得策略更好地反映最近的市场情况。

从宏观来看，市场是在一直变化的，因此针对距离交易时刻更近一些的数据进行特征的挖掘和策略的开发，就能更好地保证特征的存留以及策略的及时有效，是一个符合基本直觉的做法。而在交易资产或者金融市场出现转折点时，转折点前后的优化结果会根据转折本身的特性进行变化，这样的做法就可以更好地应对这种性质或环境的突然转变。同时，市场虽然不停地变化，但是这种变化本身在大多数情况下又是趋于稳定的，因此从过往一段时间内提取出的数据特征和相应的策略在一次交易周期内仍然有很大可能是值得信赖的。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

从操作的选择上来看，用来优化的时间段越长，则提取出的特征和策略就会由于数据量的增加而更加可信，而时间段越短，则提取出的特征和策略就越具有时效性，策略越灵活，也越容易快速地度过转折点前后的失效期。

本书主要讨论的是中低频的量化交易策略，由于交易频率较低，交易的及时性要求不是特别高，因此在使用量化交易策略进行每一次的交易判断之前，用来进行策略优化的时间和计算空间往往是比较充足的，最近数据的延迟也基本可以忽略，因此量化交易策略的实际使用者会更倾向于做这样的优化处理。

相较于较为简单的样本外检验，推进分析可以在回溯测试的过程中更为精准地反映出这种实际的处理方式。每一次优化都推进一个周期的形式，实际上就是在回溯的环境下复制实际交易中每次判断前都优化策略的做法，这也是作者在第 3 章 3.3 节中曾经提到过的，让回溯测试环境尽量与真实交易环境保持一致的最优解决方案。正是由于这种贴近真实交易的特性，在判断策略盈利能力时，推进分析下的模拟交易情况具有更高的参考价值。

当然，同样本外检验的初衷一样，推进分析的首要目的实际上也是判断量化交易策略具有多大程度的过度拟合问题。仍然用图 6-2 来做演示，对于推进过程中的每一行而言，具体的操作都相当于一次样本外检验。如果回溯测试中所使用的量化交易策略的刻画超过了数据本身的内在特征，那么在所有样本内部分都会具有过度拟合的问题，实际的反应就是针对每一行的数据集，灰色框的样本外检验结果都将大概率的劣于白色框的优化结果。当一个特征在每一行数据的检验过程中发生的概率都较大时，将多行结果合并到一起就可能会出现统计上的明显表现。因此，假如量化交易策略并不具备挖掘数据内在特性的能力，也就是不具备真实的盈利能力或预测能力，那么综合所有灰色框一起考虑时，整体结果就很可能会缺乏显著的盈利性或预测准确性。从上面的叙述中可以发现，推进分析的检验过度拟合和判断策略实际盈利能力这两个功能，在一定程度上是相互融合的。

同时，推进分析可以在回溯测试中较为有效地排除未来信息的存在。依然采用图 6-2 的例子进行说明，在每一行的数据集中，都是首先针对白色框的样本内数据进行优化，然后将优化后的具体交易策略代入与样本内数据无交集的样本外数据进行检验。这种回溯测试的方法完全不同于第 4 章和第 5 章中作者介绍择时

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

策略、选股策略时所采用的简单框架，数据的优化和检验在时间轴上是分开而且具有严格先后顺序的，因此也就在很大程度上规避了回溯测试中存在的未来信息问题。总体来看，推进分析这一回溯测试工具比较全面地照顾到了第3章中作者介绍的几个需要注意的事项，是一个在研究量化交易策略时相对而言较为稳妥的解决方案。

在实际的量化交易策略研发中，有时候会由于使用的策略方法计算量过大而导致完全按照推进分析的流程进行检验出现困难。一般而言，样本内优化需要的计算量远远大于样本外检验的计算量，因此当出现计算量过大、计算机处理水平不足的问题时，可以通过降低最优化处理的频率来有效减小总体的计算量。图6-3基于图6-2中的示例，给出了一个经过改动的推进分析的例子用以说明。在这个例子中，仍然先基于周期1到周期5的数据进行优化，但在得到参数确定的量化交易策略之后，将其放置于周期6和周期7的数据上进行检验。也就是说，在周期7上运行的策略同样也是经过周期1到周期5数据优化得到交易策略，那么图6-2中前两行的计算过程就简化为图6-3中的第一行的计算，计算量几乎减少了一半。在图6-3中，第一行的数据处理完成之后，第二行中用来优化策略的是周期3到周期7的数据，样本外检验则包括周期8和周期9的数据，以此类推，每一次都推进两个周期。因此整体而言，计算量也就几乎减少到图6-2中推进分析的一半。

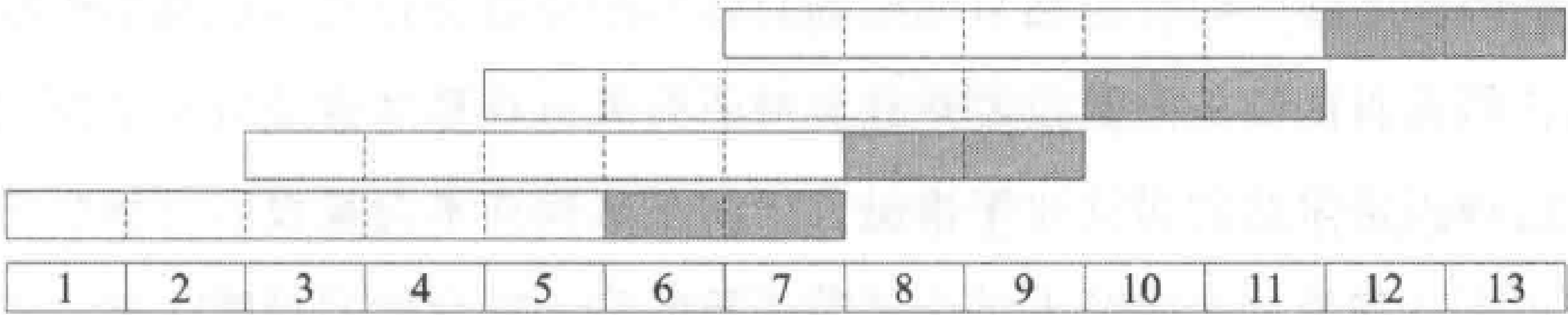


图6-3 计算紧张时改动的推进分析示例

当然，这样的改动在节省计算量的同时，也会随之生成一些其他问题。首先，这样的回溯测试过程不能够再精细地体现出实际交易下不断优化更新策略的操作过程。当然，对于在实际交易中每隔一段时间优化一次策略的使用者而言，这样的改动反而能够更好地刻画现实情况，如何选择需要根据实际情况进行判定。最重要的是，在假设市场稳定变化的前提下，这样合并检验数据的做法会使

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

得策略的有效性受到损害。

以图 6-3 为例而言，相较于图 6-2 中的标准推进分析过程，使用周期 1 到周期 5 的数据优化得到的量化交易策略，会由于交易资产和市场数据的变化而对周期 7 的数据适应能力变差，周期 9、11、13 下的检验情况也类似。因此，这样的改动虽然会降低量化交易策略回溯测试的计算量，但是也会同时增加判断策略真实能力的难度，在计算能力可以支撑的情况下作者并不推荐这种做法。

另一方面，如果实际交易也是间隔一段时间优化一次策略，那么对于相隔较远的交易决策而言，策略的适应能力同样会降低，进而弱化盈利能力。鉴于在真实交易环境下优化操作只需要进行一次，因此硬件的计算能力往往并不是策略运行的制约因素之一，在这种情况下，作者其实更不推荐在实际操作中降低中低频量化交易策略优化的频率。对于作者的这一观点，读者可以自行进行判断。

6.2 多层推进分析

推进分析通过先优化策略参数，然后将量化交易策略置于其后的数据进行检验的方法，可以较为有效地对策略所具有的过度拟合问题进行判断。但是值得注意的是，在进行推进分析的过程中，研究人员有时仍然会将一些参数固定从而方便研究。用第 4 章提到的自回归策略举例而言，当研究人员使用自回归模型预测下一个周期的收益率时，虽然自回归的具体回归系数会在逐次优化中不断地被估计出来，但是自回归的阶数设置往往会被研究人员在整个推进分析中固定为一个常数。这种固定阶数的做法对于推进分析的某次操作本身是没有任何影响的，但是问题在于，如果一个研究人员在试验了阶数为 5 的自回归择时策略，发现其推进分析的结果没有显著盈利能力之后，又试验了阶数为 6、7、8、9、10 的自回归择时策略，最后发现 9 阶自回归模型构建的择时策略具有一定的盈利能力，并将其确定为实盘使用的交易策略，就会产生过度拟合检验不完全的情况。在这个例子中，阶数实际上可以被看作策略的一个具体参数，对不同阶数的自回归策略都进行测试并选择结果最优的一种，也就可以被视为一次最优化的过程。但是直到决定使用 9 阶自回归策略进行实盘操作为止，研究者都没有针对这一参数分析

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

其过度拟合的情况。

在大多数模型优化方式比较固定的情况下，这种模型设置类的参数都是难以和模型系数一起进行优化的，因而常常会被研究人员在推进分析中设置为某一固定值。例如在上面所举的例子中，自回归模型的回归系数估计一般都会使用普通最小二乘法等计量经济学中常用的估计方法来完成，这些方法本身是不能用来同时优化自回归模型的阶数设置的。更为明显一些的例子，是采用诸如人工神经网络这样的机器学习算法进行构建的量化交易策略，这些算法的拟合技巧往往都较为复杂，而且在长时间的科学研究过程中会固化出一些公认的有效处理方法，甚至包括诸如模型预训练这样的复杂操作，不易变动。因此在绝大多数情况下，试图更改这些拟合方法进而同时优化人工神经网络隐层节点数之类的参数都是不切实际的。其他一些情况包括支持向量机中的核函数选择等，虽然不再是一个直接的参数，但是仍然能够通过人为控制来改变整个策略的实际效能，因此尝试多种核函数的做法同样等同于最优化的过程。或者更为隐蔽的训练样本大小，也应该被看作一种参数，只不过这种参数来自对推进分析本身的设置。以图 6-2 所展示的推进分析为例，如果研究者分别在白色框长度为 3、4、5、6、7 等设定的推进分析框架下测试了同一种量化交易策略，最后发现长度为 5 时策略收益最高，并在实际运行中也将优化期固定为 5 个周期，那么就同样缺乏对于白色框长度这一设置的过度拟合程度判断。

这一节中将要介绍的多层推进分析，正是为了缓解这一问题而针对推进分析流程进行的改进。多层推进分析的原理叙述起来较为简单，就是在推进分析的基础之上，针对进行交易的灰色框序列再做新的推进分析，从而研究上一层推进分析中无法检验的参数过度拟合情况。但是实际上，在量化交易策略研究中，多层推进分析的逻辑其实较为复杂，如果不小心处理就有可能会出现细节上的错误。在研究人员较难驾驭这种复杂逻辑的情况下，作者建议最好避免使用这种方法，为了测试一些参数的过度拟合程度而导致整个回溯测试结果的不可信，其实是得不偿失的。当然，接下来依然要介绍这种方法的具体操作流程，读者可以在量化交易策略的研发过程中基于现实情况来判断使用与否。

图 6-4 在图 6-2 所展示的推进分析的基础上，给出了一个两层推进分析的示

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

例用以说明。在图 6-2 中，每一个灰色框都是所在行的样本外数据，而在图 6-4 中，一部分的灰色框除了作为第一层推进分析的样本外数据以外，同时还是第二层推进分析的样本内数据，用以优化在第一层推进分析中无法同时优化的参数。以第二层推进分析中的第一行为例，在第二层参数固定为某定值的情况下，如同图 6-2 中所做的一样，基于周期 1 到周期 5 的数据进行优化得出一个最优策略，应用在周期 6 的数据上就可以得到该交易策略在周期 6 的收益率，同理，周期 7 可以根据周期 2 到周期 6 的数据优化出一个最优策略并得到相应的收益率，依此类推直到周期 10 为止。综合周期 6 到周期 10 的收益率结果，就可以得到第二层参数设为该定值时这 5 个周期上的策略收益情况，之后更换第二层参数的设置，重复上述过程，得到其他参数情况下的策略收益率。在获得第二层参数所有可能取值下的策略收益率后，就可以选出收益情况最好的一个或一组参数，这一轮操作可以看作是第二层推进分析中的样本内优化过程，如图 6-4 中的浅灰色部分所示。

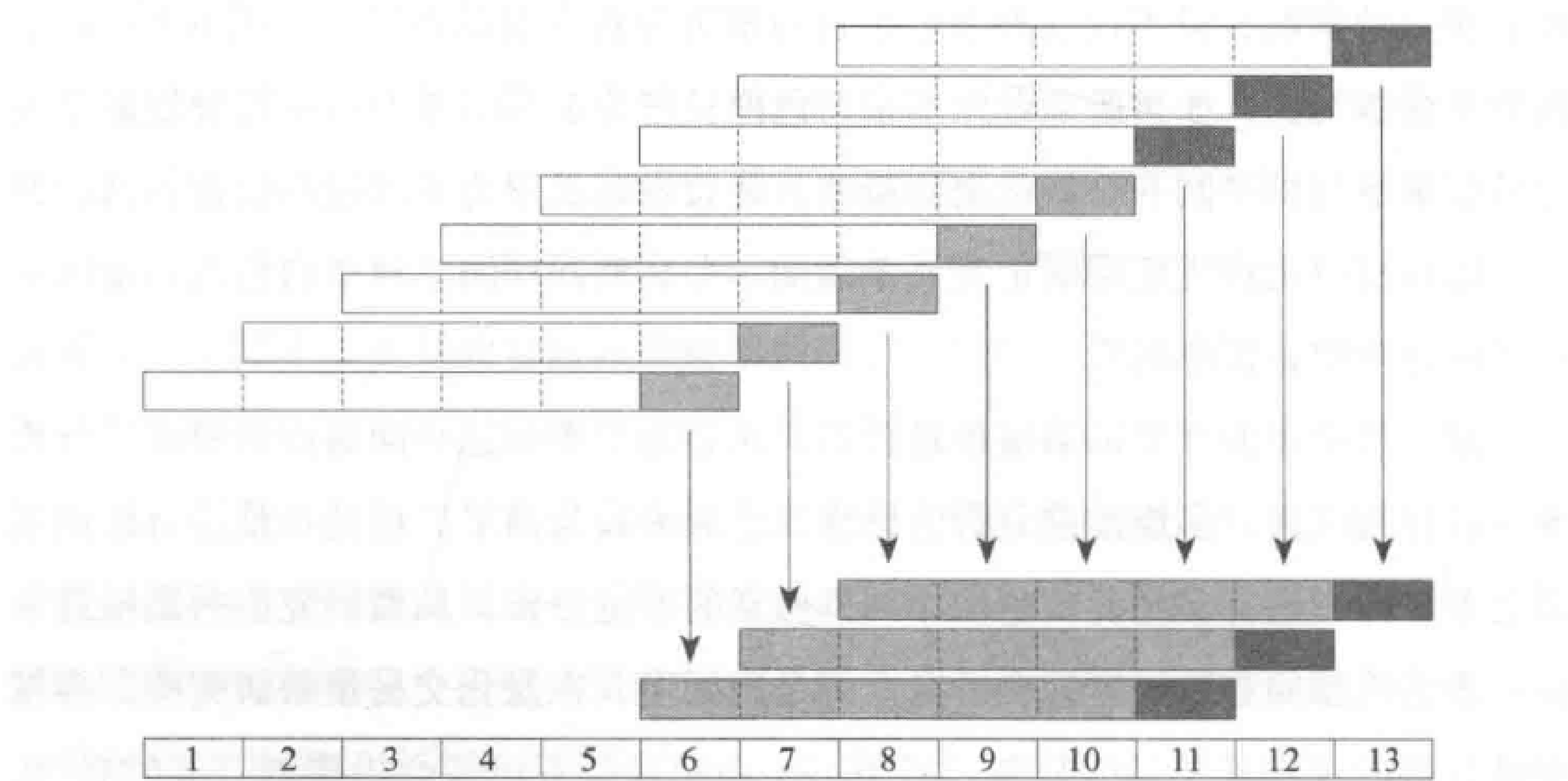


图 6-4 多层推进分析示例

在得到第二层参数的优化值之后，将策略中相应的设置固定为该优化值，然后回到第一层推进分析当中，基于周期 6 到周期 10 的数据，对第一层推进分析的参数进行优化并得到优化结果。这样，针对周期 11 而言，使用前期数据同时得到了最优的第一层参数值和第二层参数值，将基于这些参数设置的策略应用到

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

周期 11 的数据之上，就可以得到一个两层推进分析的样本外交易结果，如图 6-4 中的深灰色框所示。至此，针对第二层推进分析中第一行的操作全部完成。

然后将所有的时间窗向后推进一个周期进行同样的工作，可以完成第二行的操作。通过不断地向后推进一个周期直到整个数据样本结束，就可以得到多行的第二层推进分析结果，当然图 6-4 受空间所限只展示了三行。将深灰色框下最优策略的模拟交易结果进行综合，就可以得到该量化交易策略在两层推进分析下的策略净值走势和收益情况，从而判断出第二层参数的过度拟合程度，并更好地对整个策略的真实盈利能力加以判断。

上面介绍的是双层推进分析的具体流程，在第二层推进分析的基础之上，可以采用相似的方法进一步划分数据，进而形成第三层的推进分析。只要数据量足以应对复杂化的推进分析过程，这样的分层就可以持续地进行下去。但是需要注意的是，多层推进分析虽然可以解决一层推进分析无法完全解决的参数优化与过度拟合判断的问题，但是同时会引入新的参数设置，增加整个研发流程的复杂程度。

举例而言，当一层的推进分析已经可以在样本内优化所有的策略参数和设置时，这个时候需要固定的参数只有推进分析本身的训练样本大小，即图 6-2 中的白色框长度。此时如果使用双层推进分析来判断训练样本量的设置是否存在过度拟合问题，那么检验的也只是第一层推进分析中的训练样本量参数，而第二层推进分析又会引入一个新的训练样本量设置，即图 6-4 中的浅灰色框长度，同时优化白色框长度时的取值范围也是一个新引入的设置，这些参数、设置依然是人为设定的，并有可能存在未来信息和过度拟合的双重问题。这样看来，不管多加多少层的推进分析，人为设置的参数总是存在的，多层推进分析不能从根本上解决这个问题。

不过即使是在这种情况下，多层推进分析的引入依然是有一定作用的。在交易资产特性、市场环境缓慢变化的前提假设下，多层推进分析中更高层的设定参数应该比低层的参数变化更加稳定，因此增加推进分析层数并对上一层推进分析的设置进行过度拟合判断是一个相对而言更为稳定、可信性更高的回溯测试流程。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

但是与此同时，由于多层推进分析提高了整个回溯测试的复杂程度，因此也同时提高了对样本数据量的需求，层数越多，需要的样本数据就越多。而遗憾的是，中国的金融市场与资本市场由于历史较短，因此历史数据量相对而言是较少的，同时市场环境、政策法规的频繁变化使得可用的样本数据更为有限，在本书主要讨论的中低频量化交易策略的背景下，多层推进分析往往缺乏实际的使用基础。

因此，本节所介绍的方法虽然有其意义所在，但是实际使用需要结合具体情况认真分析，作者自己在应用时也会相当的谨慎。本节内容更重要的意义，其实是更进一步地挖掘推进分析框架的内涵、特性及其与现实交易环境的关联性，促使读者深入地理解推进分析。

6.3 推进分析下的验证

上两节介绍了推进分析和多层推进分析的操作流程，不管是简单的推进分析，还是逻辑更为复杂一些的多层推进分析，其主要目的都在于判断策略参数的过度拟合程度，以及在一个贴近真实交易的环境下模拟出策略的运行情况。在实际的量化交易策略的使用过程中，通过研究推进分析或多层推进分析得到策略的净值走势和收益情况，就可以基于一定的收益和风险标准决定是否使用该量化交易策略。

但是由于推进分析这种方法的主要目的还包括参数过度拟合程度的判断，因此在流程设计上会有所限制，基于推进分析结果直接判断实盘与否的做法在逻辑上也显得不是那么完整。本节就在推进分析的基础上，另外介绍一种用来判断量化交易策略是否进入实际使用的方法，作为前面内容的补充。

实际上，这种方法也是根据样本外检验的原理发展而成的。与前面所介绍的多层推进分析相比，该方法虽然在技术处理上十分相似，但是在逻辑上更偏向于对交易策略进行是与否的判断，因此作者称其为推进分析下的验证。

简单来说，就是将推进分析的数据分为两个部分，一个部分进行推进分析的研究，如果结果具有合格的盈利能力，则针对同样的策略在另一部分数据上进行

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

推进分析研究。在实际研究中，后一步的推进分析只进行一次，如果结果显示盈利能力依然存在，则认定该量化交易策略可以实际使用，如果策略不再具有盈利能力，则直接判断策略无用，需要更换策略思路、重新寻找其他策略。需要注意的是，这里作者说的是“更换策略思路”、而不是“更换策略设置”，因此后一次的推进分析实际上是一锤定音的决定性流程。

图 6-5 基于图 6-2 给出的推进分析示例，展示了一个推进分析下的验证具体流程，方便读者更直观地理解推进分析下的验证这一方法。图 6-5 中的横线将原本的推进分析进行了划分，针对处在横线下方的前 5 行数据，挑选多种不同的策略，各自进行普通的推进分析研究，直到找到一个在这部分样本内具有盈利能力的量化交易策略。然后将该策略置于横线上方的后 3 行数据上，进行一次普通的推进分析，查看深灰色框内的收益情况和策略净值表现，如果仍然具有盈利能力则使用该策略进行实际交易，如果策略失效则否定该策略，重新寻找其他策略。

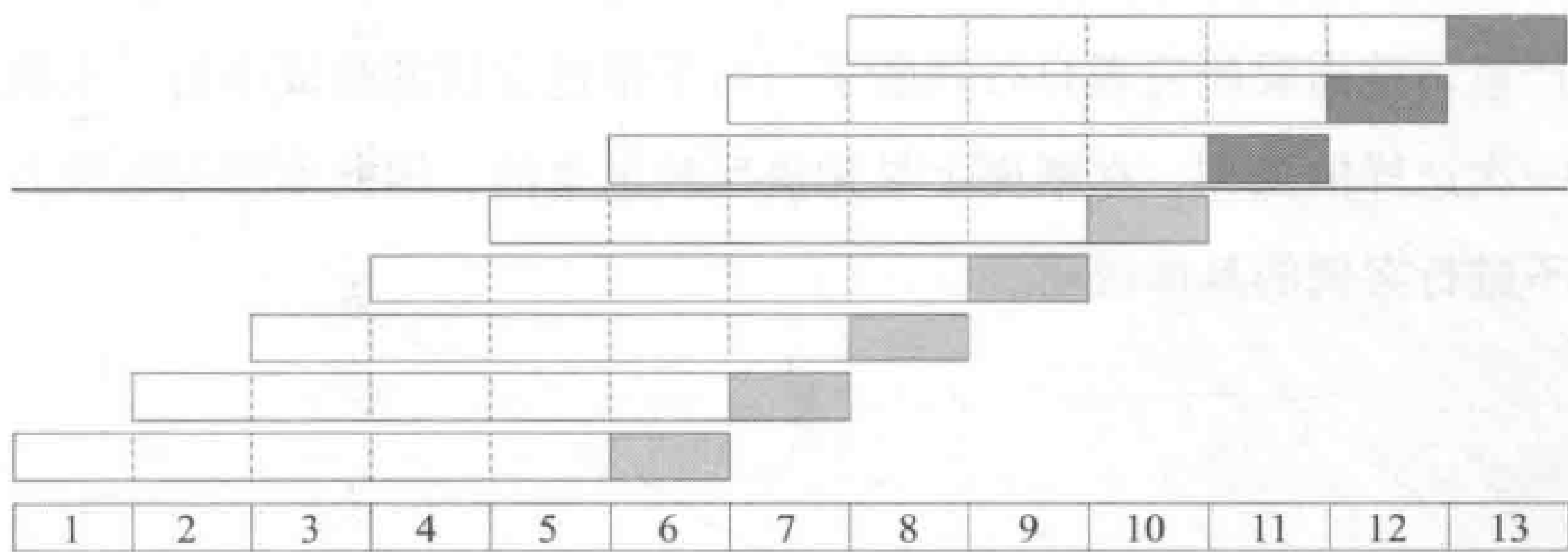


图 6-5 推进分析下的验证示例

通过如上的示例解释，读者应该可以发现，推进分析下的验证实际上就是一个样本外检验的过程，只不过前面提到的样本外检验都是针对策略模型的具体参数，而这一种样本外检验针对的则是量化交易策略本身。

从对图 6-5 中示例的具体讲解中也可以发现，该方法下样本内外的划分其实可以看作对灰色框的划分，也就是划分推进分析中模拟交易的样本外数据部分，在图 6-5 中分别用浅灰色框和深灰色框进行了标示。因此，多层推进分析也同样可以使用这种验证方法来判定策略是否进入实际应用，只需在理解多层推进分析逻辑的基础上，针对最后一层的数据样本进行样本内外的划分即可。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

但是正如多层推进分析所存在的问题一样，使用推进分析下的验证同样需要足够数据量的支持。由于中国市场上能够获取的有效数据较为有限，因此研究人员在使用推进分析下的验证前应该首先确认样本量是否充足。

更重要的问题在于，作为验证环节的后一次推进分析对策略去留具有决定性的影响，一个实际上有盈利能力的量化交易策略由于在样本数据上无法通过验证从而被淘汰的情况是有可能发生的。在推进分析下的验证中，样本内和样本外的长度设定、样本数据的具体选择等都可能会影响验证的结果，偏差较大时就会导致有效策略被错误的排除。而推进分析的验证要起到实际的决定效果，后一次的推进分析就必须严格地只进行一次，否则又会流于一种优化过程而得不到真正的检验。在这种情况下，如果由于人为设置和实际数据特性的影响导致错误地排除了有效策略，是无法通过其他办法进行补救的，研究人员必须承受这种错误发生的可能。

当然，是有效策略丢失的问题更严重，还是无效策略被应用到实际交易中的问题更严重，就需要研究者自行判断了。由于推进分析的验证中后一次推进分析只进行一次这样的操作，在纸面上是无法反映出来的，因此本书对这种方法只做介绍，不进行案例的具体说明。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

第 7 章

推进的择时策略

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

7.1 均线趋势策略的推进分析

第 4 章 4.2 节介绍了均线趋势策略的基本运行模式，并提供了一个简单的回溯测试案例。在这一节中，作者采用大体上相同的设置进行了均线趋势策略的推进分析，用来模拟实际交易中策略的运转流程，从而得到一个贴合实际环境的策略盈利能力结果，同时判断均线趋势策略两个均线参数的过度拟合程度。

双均线趋势策略的具体定义和第 4 章 4.2 节中保持一致。在交易决策的时点需要根据已知数据计算短期均线和长期均线两个值。当短期均线高于长期均线时，判断交易决策时刻的趋势为上涨趋势，按照趋势会延续的思想，认为后市会继续上涨，因此看多。而当短期均线低于长期均线时，判断交易决策时刻的趋势为下跌趋势，后市继续下跌，因此看空。具体请参见图 4-5 和相关叙述。

数据方面也和前面所选取的数据一致，交易资产是以中证指数公司编制的沪深 300 指数为标的物的沪深 300 股指期货主力合约，收集得到 2010 年 4 月 19 日至 2015 年 6 月 30 日的日收盘价数据。虽然股指期货可以在最低交易保证金为 8% 的情况下进行杠杆交易，但是在本节的回溯测试环境下，依然将杠杆率设定为 100% 进行研究。

图 7-1 展示了均线趋势策略的推进分析在这段时间的样本下是如何进行的。2010 年 4 月 19 日到 2010 年 12 月 30 日的日收盘价数据作为计算短期均线和长期均线的备用数据，数据长度为 172 日，也就是说只要长期均线的长度在 173 日以下，备用数据都可以支持整个推进分析的计算顺利完成。但是这部分数据的使用在图 7-1 中没有具体显示，希望读者注意。

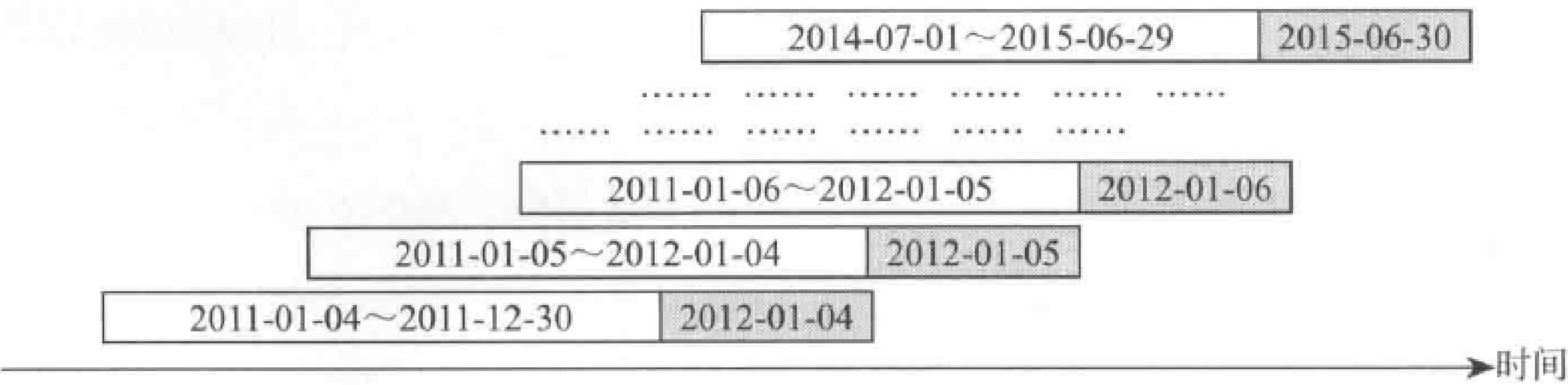


图 7-1 均线趋势策略的推进分析流程

在推进分析的第一行操作中，白色框为优化期，从 2010 年 12 月 31 日的收

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

盘开始，直到 2011 年 12 月 30 日的收盘结束，一共 244 个交易日。在这一部分样本数据下，可以通过测试不同均线组合的交易情况来寻找最优的均线趋势策略参数，具体的操作其实完全等同于第 4 章 4.2 节中所做的优化过程。由于将所有参数的可能组合都测试一遍的做法存在冗余且具有偏向性，因此优化过程在进行网格搜索时，依然使用斐波那契数列来处理均线的搜索范围，也就是使用 1 日均线、2 日均线、3 日均线、5 日均线……144 日均线一共 11 种均线的长短组合进行搜索，寻找盈利能力最好的组合，从而确定最优的均线趋势策略形式。同样限定短期均线最大搜索到 34 日均线，长期均线从短期均线的后一根开始一直到 144 日均线结束。在这一部分中，具有一致性的均线趋势策略实际上也等同于“短期均线从下向上穿过长期均线则做多，短期均线从上向下穿过长期均线则做空”。

第一行的优化操作结束之后，则按照最优结果固定短期和长期均线的长度，并用来计算 2011 年 12 月 30 日收盘时刻下的短期均线值和长期均线值。当短期均线值大于长期均线值时，就对 2011 年 12 月 30 日收盘到 2012 年 1 月 4 日收盘的走势看多，设定仓位为多头；当短期均线值小于等于长期均线值时，就对 2011 年 12 月 30 日收盘到 2012 年 1 月 4 日收盘的走势看空，设定仓位为空头。

在进行完一行的操作之后，将白色框和灰色框都分别向后推进一日，重复这种样本外检验的过程，直到最后一行为止。前面介绍推进分析时已经说明过，这个时间周期应该以判断交易的间隔为准，在本案例中使用的是日度收盘价数据，因此交易判断的最短间隔也就是一日。

最后一行中，用来优化参数的白色框从 2014 年 6 月 30 日收盘开始到 2015 年 6 月 29 日收盘结束，一共 244 个交易日，与第一行中白色框的长度一致。虽然在样本期间某些年份的总交易日数可能不是 244 日，但是按照逐行推进一日的操作流程，所有白色框的长度都应该保持为 244。最后一行中被检验的灰色框为 2015 年 6 月 29 日收盘到 2015 年 6 月 30 日收盘，也就是触及样本数据的最后一个数据。

总的来看，在推进分析的过程中，模拟买卖的行为从 2011 年 12 月 30 日收盘时开始到 2015 年 6 月 30 日收盘时结束，一共交易三年半的时间，也就是所有

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

灰色框所覆盖的区间。

需要说明的是，在第 4 章 4.2 节均线趋势策略的简单优化中，策略的参数设置在四年半的模拟交易时间内始终保持不变，因此这段时间内的均线趋势策略也等同于“短期均线从下向上穿过长期均线则做多，短期均线从上向下穿过长期均线则做空”。但是这里的模拟交易情况则不尽相同。

在灰色框所代表的模拟交易期中，三年半的时间由相对独立的 845 个交易日组成，对每一个灰色框中运行的均线趋势策略而言，参数设置都由前面 244 个交易日下的优化结果决定，因此会在这三年半的期间内产生变化。同时，由于策略逐日进行判断，而我们又希望保证“股指期货交易的杠杆率为 100%”，因此约定在每个做空日之后都重新调整杠杆率为 100%。

由于灰色框中策略设置的不一致，如果再用引号内所述的交易思维来看待整个交易过程，就会产生逻辑上的错误，出现诸如策略参数变化时买卖点不一致、杠杆率不等于 100% 之类的问题。因此，研究人员需要逐日地采用分类思维来判断多空、处理仓位变换。

当然，还有其他一些处理方法可以保证推进分析下模拟交易的正常运行，本书之所以选用这一套方法，原因在于其最贴合前面给出的量化交易策略框架。但是使用分类变换的方法来处理仓位时，可能会由于策略参数的切换导致多空在短时间内来回变动，进而带来操作上的反复，这是该方法本身存在的不足之处。

在整个推进分析 845 行的过程中，均线趋势策略的参数组合一共变动了 44 次。实际上，变动次数与用来优化的白色框的长度有内在关系，优化样本的长度越短，则变动越频繁。

将整体综合起来进行计算得知，用来优化的白色框内策略平均的年化收益率为 40.37%，而用来检验策略的灰色框的年化收益率则为 5.53%，两相比较，均线趋势策略的长短两个参数具有一定的过度拟合现象。

逐行的样本外检验结果则可以用来说明具体时间内参数的过度拟合情况，但是由于逐行下的结果波动很大、不易分辨，因此一般而言可以将总样本划分为若干个时间段并进行平均化研究。这里将灰色框三年半的时间按半年为单位划分为 7 个部分，每半年中的优化期年化收益率和检验期年化收益率情况如表 7-1 所示。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

当然，这里只是用一个例子来说明具体的做法。由于检验期整体的年化收益率过低，与优化期整体的年化收益率相差也较大，因此再进一步细分从而判断过度拟合的意义已经不大。

表 7-1　优化期年化收益率与检验期年化收益率对比　（%）

时间段	优化期年化收益率	检验期年化收益率
2012 上半年	32.62	−46.29
2012 下半年	24.16	29.45
2013 上半年	39.88	30.97
2013 下半年	35.59	−41.53
2014 上半年	36.45	−1.05
2014 下半年	31.22	143.58
2015 上半年	92.01	13.60

图 7-2 展示了均线趋势策略在推进分析的框架下，模拟交易过程以 1 为初始值的净值走势情况。在 2012 年年底、2013 年年底和 2014 年年底，策略的净值分别变为 0.8339、0.7297 和 1.1329。2012 年策略净值有所下跌，在 2013 年中期回归初始值附近，随后下跌至最低点。在样本期最后的一年多时间内，净值保持了一个上升态势，直到最后一段时间有一个较为明显的下跌，到 2015 年 6 月 30 日收盘时的策略净值为 1.2074。

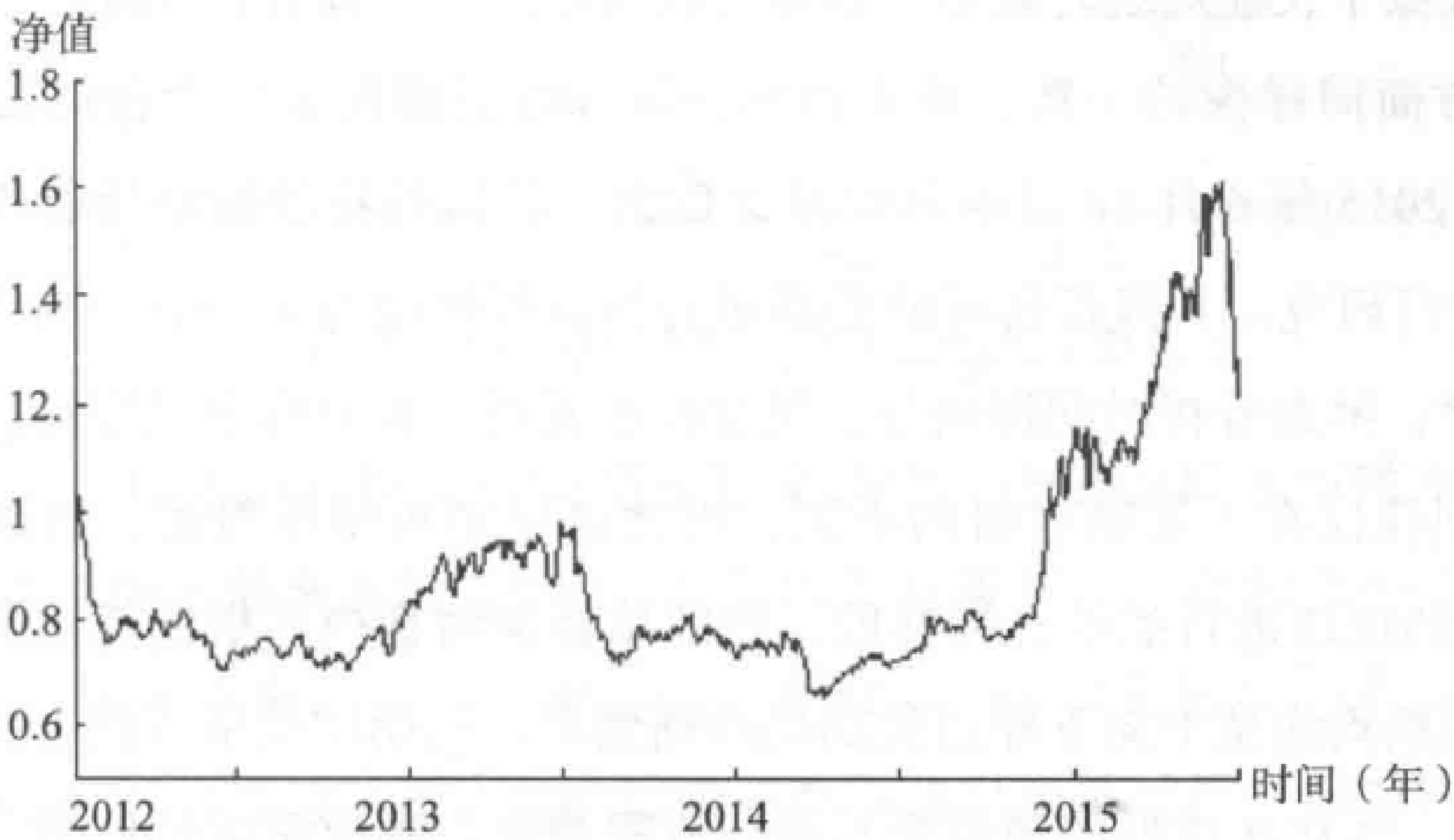


图 7-2　推进分析下均线趋势策略的净值走势

基于整个交易过程来看，三年半的策略收益率为 20.74%，年化收益率为 5.53%，可以说缺乏较为明显的盈利能力。同期对比来看，均线趋势策略的简单

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

优化下，2011 年 12 月 30 日收盘到 2015 年 6 月 30 日收盘时的收益率为 95.54%，是这里 20.74% 的将近 5 倍。

在整个三年半的均线趋势策略模拟交易当中，净值的最大回撤率为 36.71%，回撤的发生从 2012 年 1 月 5 日收盘的 1.0283 开始，一直到 2014 年 4 月 9 日收盘时的 0.6508 结束，整个回撤过程一共经历 543 个交易日，是一个相对而言较长的回撤时间。经过计算，均线趋势策略在整个交易过程中平均的年度最大回撤率为 20.25%，回撤水平也是比较大的。

7.2 均线反转策略的推进分析

第 4 章 4.3 节介绍了均线反转策略的基本运行模式，并提供了一个简单的回溯测试案例。在这一节中，作者同样采用与前面大体相同的设置对均线反转策略进行推进分析，模拟实际交易中策略的运转流程，进而得到一个贴合实际环境的策略盈利能力结果，同时判断反转趋势策略两个均线参数的过度拟合程度。双均线反转策略的具体定义和第 4 章 4.3 节中保持一致，在交易决策的时点需要根据已知数据计算短期均线和长期均线两个值，当短期均线高于长期均线时看空，而当短期均线低于长期均线时看多。具体可以参见图 4-7 和相关叙述。

数据方面同样保持一致，收集得到沪深 300 股指期货主力合约从 2010 年 4 月 19 日到 2015 年 6 月 30 日的日收盘价数据，且依然将股指期货的杠杆率设定为 100% 进行研究。均线反转策略的推进分析运行情况与上一节中的均线趋势策略如出一辙，推进分析时间段划分、优化样本长度、检验样本长度等都保持同一设置，区别仅仅在于策略逻辑的不同，因此这里不再多做赘述，请读者根据图 7-1 和相关的描述进行类推。同样的，均线反转策略在白色框的优化期中，可以等同于“短期均线从上向下穿过长期均线则做多，短期均线从下向上穿过长期均线则做空”。而在灰色框的模拟交易运行过程当中，则应该保持使用分类思维来处理仓位变换，根据长短均线的大小对比来逐日判别多空并形成交易。

在整个推进分析 845 行的过程中，均线趋势策略的参数组合一共变动了 58 次。综合计算可知，用来优化的白色框内策略平均的年化收益率为 31.85%，而

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

用来检验策略的灰色框的年化收益率则为 -10.04%，差距比较大，甚至盈亏方向都不一致。

图 7-3 展示了均线反转策略在推进分析的框架下，模拟交易过程以 1 为初始值的净值走势情况。在 2012 年年底、2013 年年底和 2014 年年底，策略的净值分别变为 1.0826、1.2278 和 0.9482。在交易样本期的前半段内，策略的净值水平大部分时间都处在 1 以上，2014 年 2 月 19 日收盘时达到最高点 1.3222。但是在 2014 年 9 月之后，策略净值经历了一个非常直接的下跌，2015 年 6 月 30 日收盘时达到 0.6904。基于整个交易过程来看，三年半的策略收益率为 -30.96%，年化收益率为 -10.04%，是一个表现较差的策略净值情况。同期对比来看，均线反转策略的简单优化下，2011 年 12 月 30 日收盘到 2015 年 6 月 30 日收盘时的收益率为 42.12%。

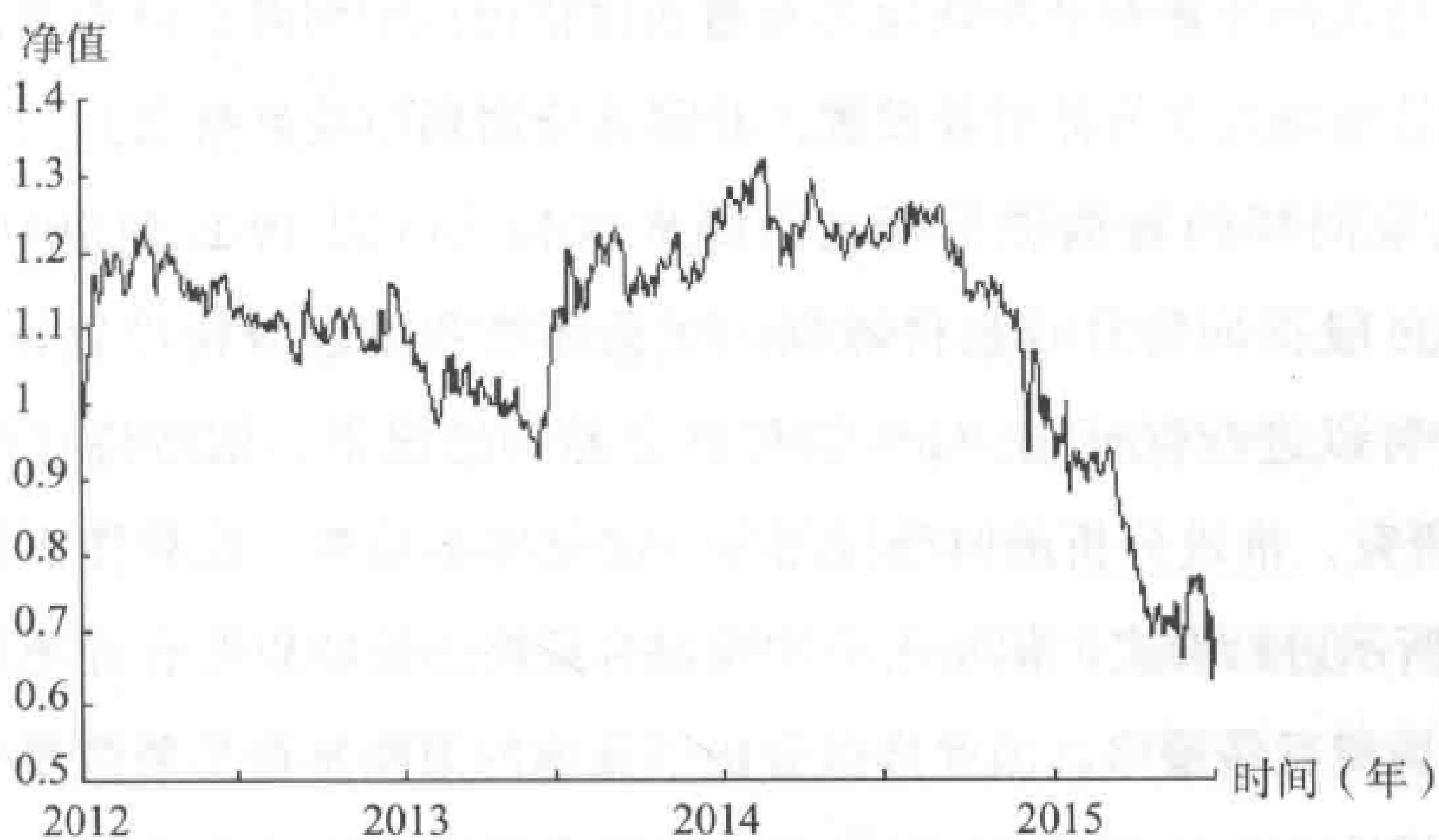


图 7-3 推进分析下均线反转策略的净值走势

在整个三年半的均线反转策略模拟交易中，净值的最大回撤率为 52.16%，回撤的发生从净值最高点的 2014 年 2 月 19 日收盘起，一直到 2015 年 6 月 29 日收盘时的 0.6325 结束，整个回撤过程一共经历 333 个交易日，其实也就是样本期的后半段时间。经过计算，均线趋势策略在整个交易过程中平均的年度最大回撤率为 21.07%，回撤水平也是比较大的。

实际上，推进分析下的均线反转策略表现较差，将其反向操作倒是能够取得一定的收益。但是在推进分析框架下，研究人员是采用前期最优的策略进行交

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

易，具有基本的交易逻辑和经济直觉，反向操作后将丧失这种基础的逻辑性，作者不推荐这样的处理行为。

7.3 均线混合策略的推进分析

在第 4 章中，作者基于均线工具分别介绍了均线趋势策略和均线反转策略这两种量化择时策略。由于在简单的优化和回溯测试框架之下，量化交易策略在整个策略样本中具有一致性，因此一旦确定策略思想是趋势或者反转，就不能在交易过程中变化了。但是在这一章中，由于推进分析中的每一行操作都相对独立，每一步推进都需要重新判断一次分类，并基于分类思维来处理仓位变换和交易操作，因此可以引入更加灵活的策略调整手段，作者称之为均线混合策略。实际上，就是将均线趋势策略和均线反转策略在优化过程中同时进行考虑，选择盈利能力最强的策略方式和具体参数设置，并将其应用到检验数据之上。

仍然选取同样的数据进行研究，即从 2010 年 4 月 19 日到 2015 年 6 月 30 日的沪深 300 股指期货日收盘价数据。交易环境和推进分析设置也保持一致，在股指期货可以进行保证金为 8% 的杠杆交易的情况下，依然将杠杆率设定为 100% 进行研究，推进分析的时间段划分、优化样本长度、检验样本长度等均依照图 7-1 中所示进行设定。但是由于均线混合策略的策略思路有异于单纯的均线趋势策略或均线反转策略，因此推进分析下具体的策略执行会稍微复杂一些，下面对此进行描述，方便读者更好地理解均线混合策略的推进分析。

在推进分析的第一行操作中，白色框为优化期，从 2010 年 12 月 31 日的收盘开始，直到 2011 年 12 月 30 日的收盘结束，一共 244 个交易日。在这一部分样本数据下，首先通过测试不同均线组合的交易情况来寻找最优的均线趋势策略参数，然后通过测试不同均线组合的交易情况来寻找最优的均线反转策略参数。这两个优化过程在进行网格搜索时，都依然使用斐波那契数列来处理均线的搜索范围，且限定短期均线最大搜索到 34 日均线，长期均线从短期均线的后一根开始一直到 144 日均线结束。在这 244 个交易日的优化样本期上，如果最优均线趋势策略的收益高于最优均线反转策略的收益，那么选取最优参数下的均线趋势策

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

略作为最终的优化结果，反之则选取最优参数下的均线反转策略作为最终的优化结果。

在得到白色框的优化结果之后，就按照最优结果固定短期和长期均线的长度，并用来计算 2011 年 12 月 30 日收盘时刻下的短期均线值和长期均线值。基于最终选取的均线趋势策略或均线反转策略，通过比较短期均线值和长期均线值的大小，就可以得到对 2011 年 12 月 30 日收盘到 2012 年 1 月 4 日收盘走势的多空看法，进而设定当前仓位。

在完成第一行的操作之后，将白色框和灰色框都分别向后推进一日，重复同样的样本外检验过程，然后持续推进直到最后一行为止。总的来看，在推进分析的过程中，模拟买卖的行为从 2011 年 12 月 30 日收盘时开始，到 2015 年 6 月 30 日收盘时结束，一共交易三年半的时间，也就是所有灰色框所覆盖的区间。由于逐行进行优化，灰色框中的策略设置同样也是不一致的，而且趋势和反转会交替出现。因此研究人员依然需要采用分类思维来处理仓位变换。当然，也正是因为推进分析的框架和这种分类思维的使用，使得在推进分析中可以采用更加灵活的组合策略，以及回溯研究均线混合策略这种在第 4 章中无法研究的策略。这也是推进分析更贴近于实际交易环境的一种具体体现，在实际交易中，交易人员是可以根据历史数据上的研究自由选择合适的策略的。

在整个推进分析 845 行的过程中，均线混合策略的参数组合一共变动了 45 次，介于均线趋势策略和均线反转策略的变动次数之间。综合计算可知，用来优化的白色框内策略平均的年化收益率为 55.63%，而用来检验策略的灰色框的年化收益率则为 8.77%，这两个结果都比前两节的相关结果要高。就这两者而言，优化样本下的收益率更高是一个非常容易理解的结果，原因在于优化过程综合选择了均线趋势策略和均线反转策略这两个策略中更优异的一个，因此必然会比两个策略分开讨论的情况都要好。而检验期下的收益率水平有所提升，就是一个对于量化交易策略而言具有实际意义的结果了，其说明至少在回溯测试的样本下，策略的混合改进确实额外带来了一定的盈利能力。

图 7-4 展示了均线混合策略在推进分析的框架下，模拟交易过程以 1 为初始值的净值走势情况。在 2012 年年底、2013 年年底和 2014 年年底，策略的净值

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

分别变为 0.7852、0.9101 和 1.2590。在 2012 年年初，净值就迅速下降到 0.8 左右，并一直较为稳定的运行直到 2014 年 11 月，随后的策略净值水平有了一个比较大的提升。在最后时期策略净值又有一定的下降，2015 年 6 月 30 日收盘时的净值为 1.3419。基于整个模拟交易过程来看，三年半期间当中均线混合策略的总收益率为 34.19%，而净值的最大回撤率为 28.38%，回撤的发生从 2012 年 1 月 5 日收盘的 1.0283 开始，一直到 2013 年 2 月 26 日收盘时的 0.7364 结束，整个回撤过程一共经历 274 个交易日。经过计算，均线混合策略在整个交易过程中平均的年度最大回撤率为 14.69%。

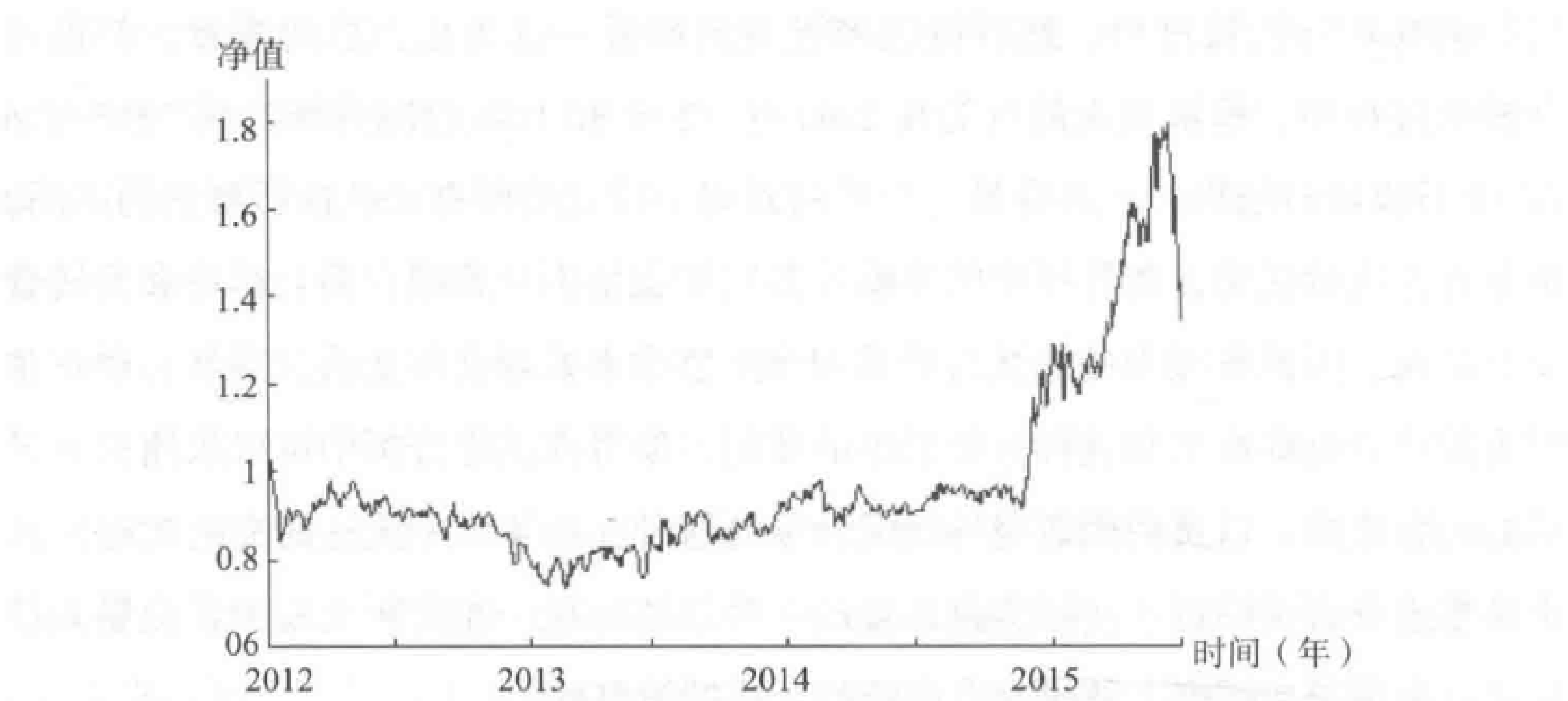


图 7-4　推进分析下均线混合策略的净值走势

虽然均线混合策略在收益率和回撤水平上都比单独的均线趋势策略和均线反转策略表现更好，但是年化收益率仍然只有 8.77%，而模拟交易中的平均年度最大回撤率为 14.69%。同时，在三年半的交易检验期中，净值在相当长的一段时间内低于初始净值且没有明显变动。实际上，综合来看，三种均线策略的推进分析结果均没有很好的表现。通过这些研究结果来判断，这三种均线策略均不适合应用于实际交易中。

当然，作为趋势类交易策略和反趋势类交易策略的典型代表，均线构成的量化择时策略是否存在实际的盈利能力显然不是一个简单的推进分析结果可以决定的，样本数据的更换、推进分析具体设置的变化等都可能改变最终的结果和结论，对趋势、反转的交易逻辑抱有信心的研究者有必要通过更多的研究来验证自己的理念。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

7.4 自回归策略的推进分析

第 4 章 4.4 节介绍了自回归策略的基本运行模式，并提供了一个简单的回溯测试案例。在这一节中，作者将采用大体上相同的设置对自回归策略进行推进分析，用来模拟实际交易中策略的运转流程，从而得到一个贴合实际环境的策略盈利能力结果，同时判断自回归模型参数估计的过度拟合程度。自回归策略的具体设置与第 4 章 4.4 节中保持一致，仍然是通过前 9 日的日度收益率回归得到交易时刻的日度收益率，基于收益率预测值的情况给出多空的看法，即交易时刻的日度收益率预测值为正时看多，为负时看空。具体的模型解释和策略详情请参见图 4-9 和相关叙述。

数据方面和前几节推进分析中所选取的数据一致，收集得到沪深 300 股指期货主力合约从 2010 年 4 月 19 日到 2015 年 6 月 30 日的日收盘价数据。在本节的回溯测试环境下，依然忽略股指期货可以进行保证金为 8% 的杠杆交易，将杠杆率设定为 100% 进行研究。数据在推进分析框架下的划分也可以沿用图 7-1 中的演示，但是需要注意的是，使用均线策略时，2010 年 4 月 19 日到 2010 年 12 月 30 日的日收盘价数据作为计算短期均线和长期均线的备用数据，数据长度为 172 日，也就是说只要长期均线的长度在 173 日以下，备用数据都可以支持整个推进分析的计算顺利完成。同样的，这段时间的收盘价数据在本节中就是估计自回归模型的备用数据，因此备用数据最高能支持 172 阶的自回归模型进行参数估计。而均线策略最多测试到 144 日均线，这里的自回归仅为 9 阶，因此本节实际使用的备用数据是要远远少于前面三节的。由于图 7-1 没有具体体现备用数据这一部分，所以特别做出说明，以免读者误解。

在推进分析的第一行操作中，白色框为优化期，从 2010 年 12 月 31 日的收盘开始，直到 2011 年 12 月 30 日的收盘结束，一共 244 个交易日。这 244 个交易日的日度收益率与各自之前的 9 个日度收益率组成了 244 组数据，针对这 244 组数据进行自回归模型的估计，得到相应的参数估计结果。然后使用估计好的自回归模型，以及 2011 年 12 月 19 日收盘到 2011 年 12 月 30 日收盘之间的 9 个日度收益率数据，来预测 2011 年 12 月 30 日收盘到 2012 年 1 月 4 日收盘的

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

收益率。该收益预期为正时，就对 2011 年 12 月 30 日收盘到 2012 年 1 月 4 日收盘的走势看多，设定仓位为多头；该收益预期为负时，就对 2011 年 12 月 30 日收盘到 2012 年 1 月 4 日收盘的走势看空，设定仓位为空头。之后逐行推进并进行相似的操作，直到样本期结束，可以得到三年半的灰色框检验期下具体的模拟交易结果。

通过上面的叙述可以发现，在逐行的操作当中，三种均线策略下的推进分析都是先在优化期对策略的参数进行优化，选取盈利能力最强的策略作为优化结果，然后将优化好的策略置入检验期进行多空的判断。而在自回归策略下，优化期主要用来估计具体的模型，得到拟合程度最佳的模型参数，然后将拟合好的模型置入检验期，再根据模型形成策略判断多空。简单来说，均线策略由优化期进入检验期的是具体的策略，而自回归策略由优化期进入检验期的则是策略中的模型。这是图 4-2 和图 4-4 所代表的两类量化交易策略在进行推进分析时非常大的一个区别，同时也是用来进行实际交易时非常大的一个区别，应该加以注意。

首先来看自回归模型在逐行样本外检验中的过度拟合程度。在前面均线策略中，由于样本内和样本外具有策略上的直接联系，因此过度拟合的判断只需要直接比较策略收益情况即可。但是在自回归策略中，直接关联样本内外的是具体的自回归模型，因此过度拟合更大程度上来自收益率预测值和真实收益率之间的差距。

在计量经济学研究中，往往更倾向于将这类做法称为模型预测能力的检验。当然，过度拟合是策略的偏误，而预测能力则是模型的效能，一个负面一个正面，在逻辑上是有区别的，希望读者更多地去体会。

将整体结果综合起来进行计算得知，灰色框中收益率预测值的年化值为 6.75%，真实收益率的年化值为 19.38%。需要说明的是，这两个值越接近，则说明模型预测能力越强，真实收益率超过收益预测值过多反而是模型能力不强的表现。但是这种情况下的结果却是策略盈利能力增加，是一个模型逻辑和实际效用相悖的表现，需要具体情况具体分析。

将灰色框三年半的时间按半年为单位划分为 7 个部分，每半年中的预期收益年化值和真实收益率年化值的情况如表 7-2 所示。当然，由于预测的特殊性，两

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

者的差距其实是比较大的，研究人员可以通过其他手段来进一步判断模型的预测能力。

表 7-2 检验期的收益率预测值与真实收益率对比 （%）

时间段	收益率预测值	真实收益率
2012 上半年	1.67	9.73
2012 下半年	−4.10	5.67
2013 上半年	−3.26	−26.99
2013 下半年	1.72	16.93
2014 上半年	−1.49	−15.61
2014 下半年	26.71	178.22
2015 上半年	31.91	48.68

当然，我们也可以依照前面的处理，通过比较白色框和灰色框下策略的年化收益率，来判断策略在推进分析下的过度拟合程度。将整体综合起来进行计算得知，用来优化的白色框内策略平均的年化收益率为 72.10%，而用来检验策略的灰色框的年化收益率则为 21.53%，过度拟合现象还是比较明显的。

同样的，将灰色框三年半的时间按半年为单位划分为 7 个部分，每半年中的优化期年化收益率和检验期年化收益率情况如表 7-3 所示。

表 7-3 优化期年化收益率与检验期年化收益率对比 （%）

时间段	优化期年化收益率	检验期年化收益率
2012 上半年	67.30	14.60
2012 下半年	60.20	61.47
2013 上半年	78.11	28.24
2013 下半年	113.10	17.94
2014 上半年	70.18	−7.05
2014 下半年	57.39	80.60
2015 上半年	64.09	−16.66

图 7-5 展示了自回归策略在推进分析的框架下，模拟交易过程以 1 为初始值的净值走势情况。在 2012 年年底、2013 年年底和 2014 年年底，策略的净值分别变为 1.3603、1.6730 和 2.1676。整个模拟交易期内，策略净值存在一个较为稳定的上升趋势，只在 2014 年年末 2015 年年初出现了一个比较大的震荡，到 2015 年 6 月 30 日收盘时的策略净值为 1.9787。基于整个交易过程来看，三年半

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

的策略收益率为97.87%，折算下来的年化收益率达到21.53%。同期对比来看，自回归策略的简单优化下，2011年12月30日收盘到2015年6月30日收盘时的收益率为96.56%。推进分析下的自回归策略与存在未来信息问题的简单优化收益表现基本一致，而且净值走势具有更高的平稳性，这表明自回归策略通过了推进分析的初步筛查。

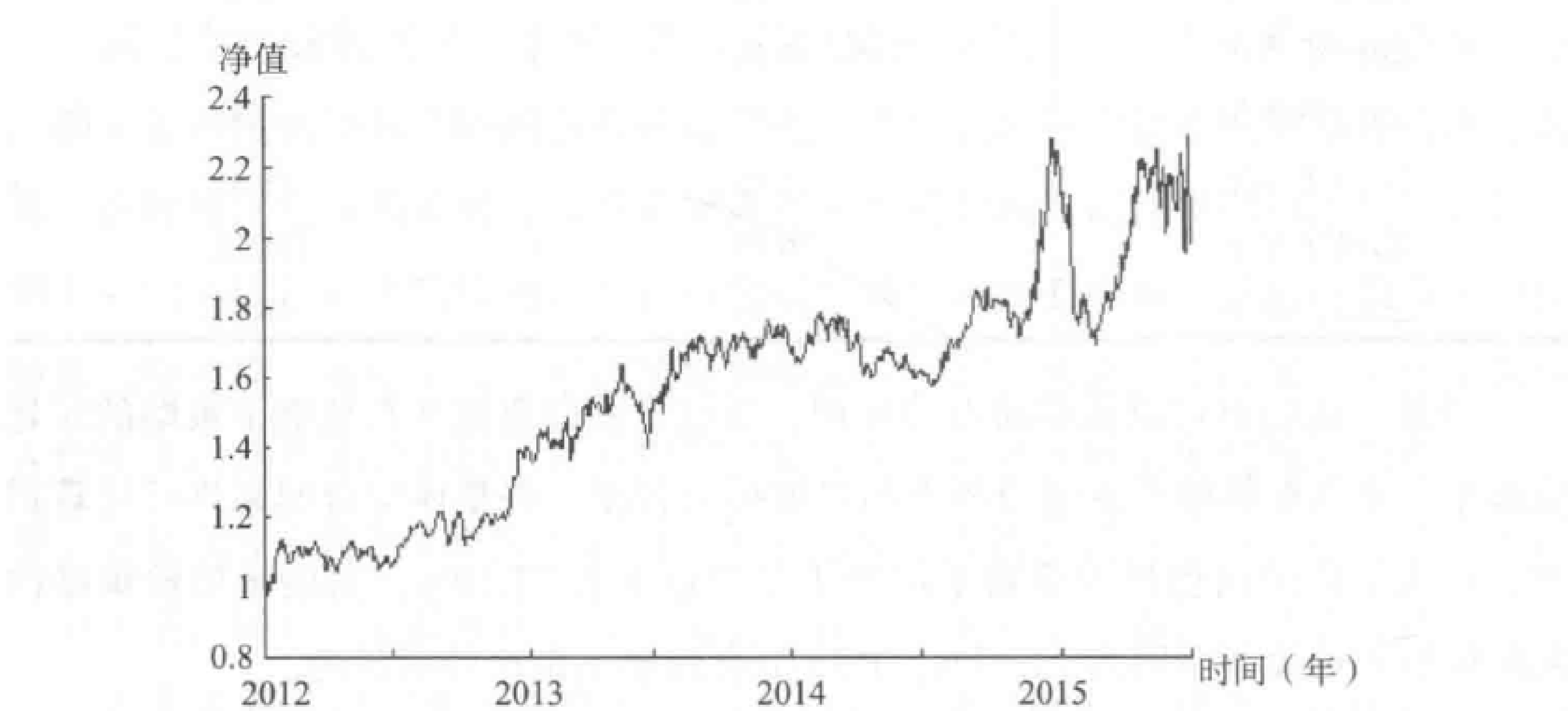


图7-5 推进分析下自回归策略的净值走势

在整个三年半的自回归策略模拟交易当中，净值的最大回撤率为25.92%，回撤的发生从2014年12月22日收盘的2.2796开始，一直到2015年2月25日收盘时的1.6887结束，整个回撤过程一共经历40个交易日，是一个相对而言较为急促的回撤过程。而且值得注意的是，这一段回撤期与因子选股策略的回撤期有非常大的重合程度，中国的股票市场和股指期货市场在这段时期出现了什么样的特性变化，使得择时和选股的多种策略同时失效，是一个值得深思的问题。此外经过计算，自回归策略在整个交易过程中平均的年度最大回撤率为14.52%，和年化收益率比较来看，这一回撤水平是可以接受的。

7.5 自回归策略的多层推进分析

上一节以沪深300股指期货为样本介绍了自回归策略的推进分析流程，并发现在特定的样本期以及策略设置下得到了一个相对积极的盈利结果。不过由于自

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

回归模型阶数这一模型设置参数没有和回归系数一起进行优化，而是固定为9阶并进行模型的估计，因此该参数就存在过度拟合验证不完全的问题。而且，作者可以直截了当地告诉读者，9阶这一设置实际上是作者在全样本上试验了若干个阶数后选取的一个表现较为优秀的情况，这里面很明显地存在未来信息使用的问题。因此在这一节中，作者将主要针对阶数这一参数设置进行自回归策略的双层推进分析，一是进一步判断自回归策略的实际盈利能力，更重要的是为读者提供一个可以参考的多层推进分析示例。自回归策略沿用上面的构造方式，这里不多做赘述，具体请参考上一节以及第4章4.4节的相关内容。

获取的数据仍然和前面一致，交易资产为沪深300股指期货的主力合约，具体包括从2010年4月19日到2015年6月30日的日收盘价数据。并且在本节的回溯测试环境下，依然将股指期货交易的杠杆率设定为100%进行研究。图7-6描述了自回归策略双层推进分析中第二层的数据划分和运行情况，而推进分析第一层的设置与上一节的做法完全保持一致，为了简化示意图而没有加入，希望读者自己联想全过程。在图7-6中，第二层计算最早的时间点为2011年12月30日收盘，因此第一层推进分析的首次优化时间为2010年12月31日收盘到2011年12月30日收盘，再之前的数据都作为备用数据加以使用，这和上一节的情况是一致的。

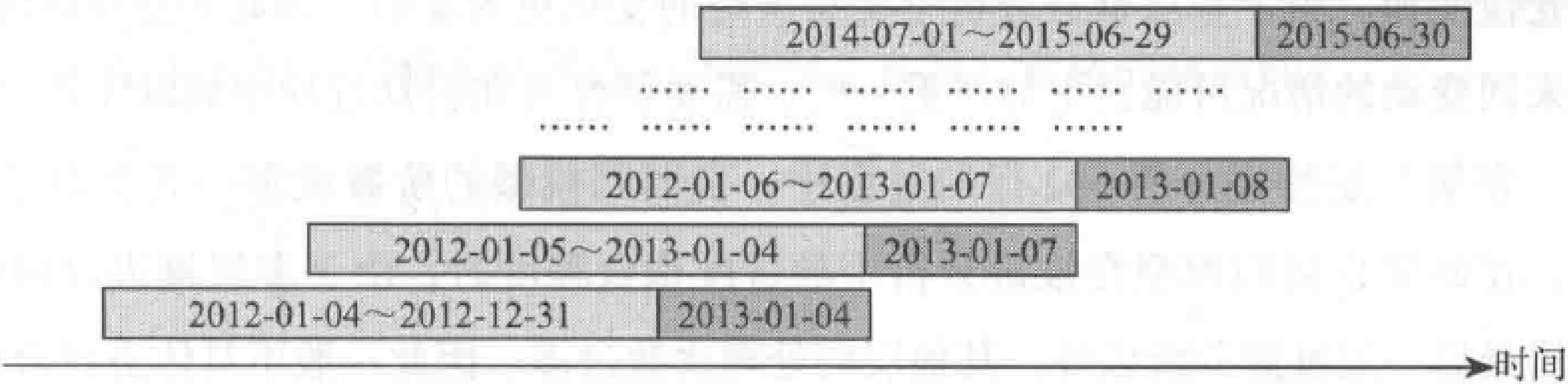


图7-6　自回归策略的多层推进分析流程

在第二层推进分析的第一行操作中，浅灰色框为阶数设置的优化期，从2011年12月30日的收盘开始，直到2012年12月31日的收盘结束，一共243个交易日。需要说明的是，在这个例子当中，第二层推进分析选择优化期长度时仍然基于一整年的考虑进行确定，但是由于2011年全年和2012年全年的交易日数量不一致，因此推进分析第一层和第二层的具体设置也就相应地存在区别，也

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

就是第一层中白色框的长度为244，第二层中浅灰色框的长度为243。而且由于推进分析逐行推进一日的操作特性，这两个设置将一直保持到推进分析流程结束，不再改变。

再回到第二层推进分析第一行的操作中来，浅灰色框中的优化主要针对自回归模型的阶数这一参数，实际上就是依次选取从1到20的阶数设置，然后分别在固定阶数的情况下，进行完全等同于上一节中自回归策略推进分析的操作，最后根据不同阶数情况下所得到模拟交易盈利结果，选取一个盈利能力最优的阶数。在确定了第一行的最优阶数之后，就可以将该阶数设置固定到深灰色框的自回归模型之上，然后进行相似的第一层推进分析操作，也就是估计出自回归模型系数之后，预测2012年12月31日收盘到2013年1月4日收盘的日度收益率，根据正负值判断多空，进而得到模拟的交易情况。

在进行完第一行的操作之后，将浅灰色框和深灰色框都分别向后推进一日，重复同样的流程，直到最后一行为止。总的来看，在双层推进分析的过程中，模拟买卖的行为从2012年12月31日收盘时开始，到2015年6月30日收盘时结束，一共交易两年半的时间，也就是所有深灰色框所覆盖的区间。与推进分析的情况一样，多层推进分析中深灰色框所应用的交易策略仍然具有设置上的不一致性，而且这种不一致性程度更高，因此研究人员同样需要逐日采用分类思维来处理仓位变换。由于多层推进分析中交易策略的变化更为复杂，因此多空在短时间内来回变动的情况可能会更加严重一些，需要研究者在研发过程中加以注意。

在第二层推进分析602行的过程中，自回归模型的阶数设置一共变动了23次。在研究自回归模型在推进分析中的过度拟合程度时，由于多层推进分析中，除了最后一层推进分析以外，其他层的处理次数过多，因此一般不具体进行分析。这里仅针对第二层推进分析来研究自回归模型阶数设置的过度拟合程度。在上一节中，由于检验期内需要根据收益预期来判断买卖，因此自回归策略过度拟合程度的判断来自收益率预测值和真实收益率之间的差距。而在这一节中，第二层推进分析的样本内和样本外已经具有策略上的直接联系，所以过度拟合的判断只需要直接比较策略收益情况即可，这和本章前三节处理均线策略时的做法反而更为接近。将整体综合起来进行计算得知，用来优化的浅灰色框内策略平均的年化收

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

益率为 32.77%，而用来检验策略的深灰色框的年化收益率则为 16.71%，相对而言两者是较为接近的。此外，将深灰色框两年半的时间按半年为单位划分为 5 个部分，每半年中的优化期年化收益率和检验期年化收益率情况如表 7-4 所示。

表 7-4 第二层的优化期年化收益率与检验期年化收益率对比 (%)

时间段	优化期年化收益率	检验期年化收益率
2013 上半年	35.95	2.31
2013 下半年	47.85	−5.03
2014 上半年	29.11	13.44
2014 下半年	24.10	87.28
2015 上半年	28.13	4.89

图 7-7 展示了自回归策略在双层推进分析的框架下，模拟交易过程以 1 为初始值的净值走势情况。在 2013 年年底和 2014 年年底，策略的净值分别变为 0.9857、1.4367。在整个模拟交易期内，策略净值依然存在一个较为明显的上升趋势，但在 2014 年年末 2015 年年初也出现了一个比较大的震荡，到 2015 年 6 月 30 日收盘时的策略净值为 1.4714。基于整个交易过程来看，两年半的策略收益率为 47.14%，折算下来的年化收益率达到 16.71%。同期进行对比，上一节中自回归策略的推进分析下，2012 年 12 月 31 日收盘到 2015 年 6 月 30 日收盘时的收益率为 45.46%，双层推进分析下的自回归策略的盈利能力还要略强一些。考虑到上一节中的阶数 9 是作者多次试验下的优选结果，本身带有未来信息的问题，双层推进分析的研究结果进一步说明了自回归策略的有效性。

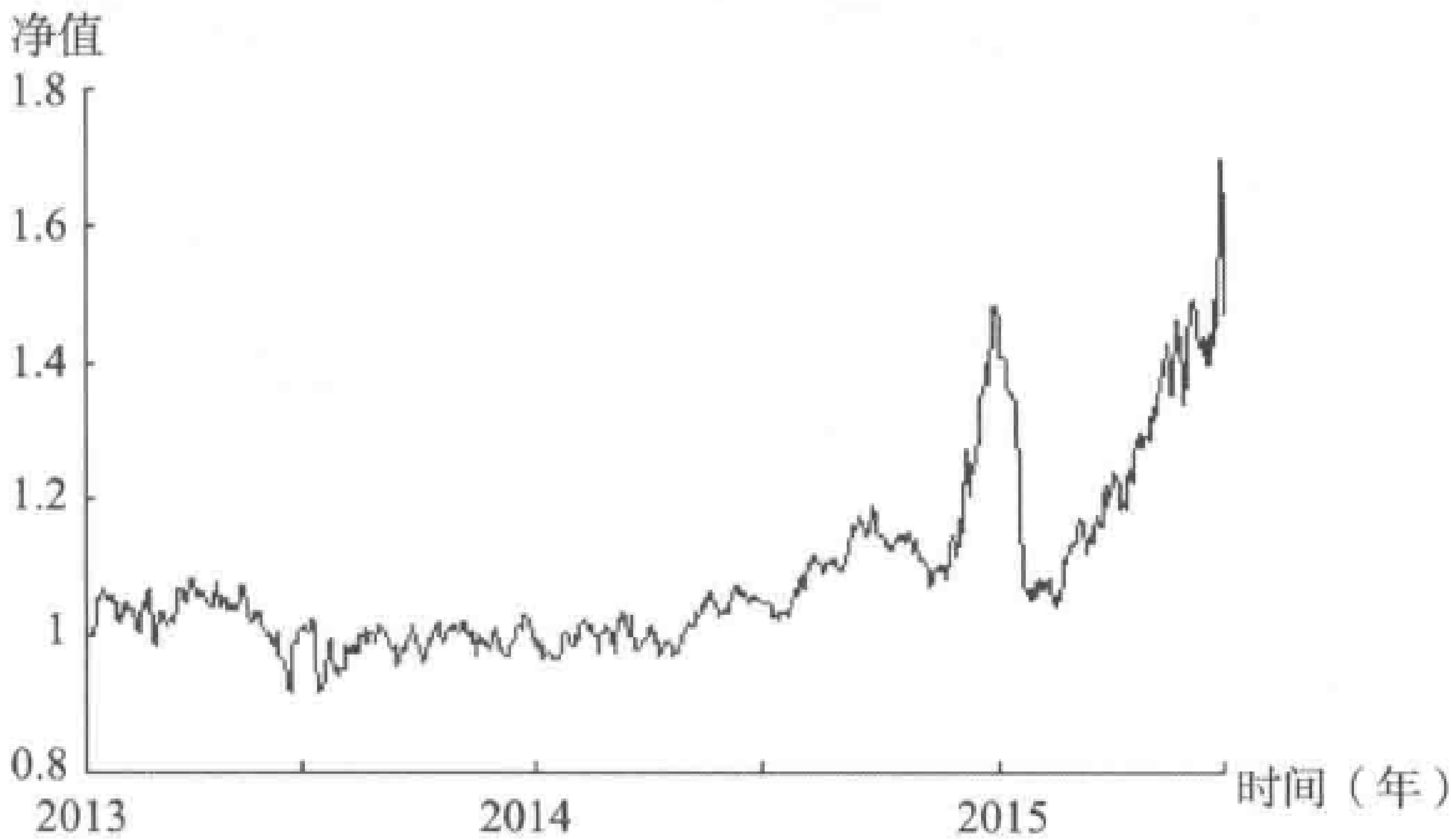


图 7-7 双层推进分析下自回归策略的净值走势

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

在整个两年半的自回归策略模拟交易当中，净值的最大回撤率为29.59%，回撤的发生从2014年12月26日收盘的1.4840开始，一直到2015年2月13日收盘时的1.0448结束，整个回撤过程一共经历33个交易日，较之上一节中的回撤期更短而且正处于上一节回撤期的中间。经过计算，自回归策略在整个交易过程中平均的年度最大回撤率为16.45%，和年化收益率比较来看，这一回撤水平依然是可以接受的。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

第8章

推进的选股策略

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

8.1 多因子选股策略的推进分析

第 5 章 5.4 节介绍了多因子选股策略的基本运行模式，并提供了一个简单的回溯测试案例。在这一节中，作者采用大体上相同的设置进行了多因子选股策略的推进分析，用来模拟实际交易中策略的运转流程，从而得到一个贴合实际环境的策略盈利能力结果，同时考察策略的过度拟合程度。多因子选股策略的具体定义和第 4 章 4.4 节中保持一致，仍然是使用估计完成的回归式计算得到所有股票收益率的预测值，然后将股票按照收益率预测值从高到低进行排序，再按顺序等分为若干个部分，每一等份中的股票又按照等权重进行组合，也就是买入各只股票占用的资金量相等的组合。交易中，选取排序中的第一个等份进行持仓操作。具体请参见第 5 章 5.4 节前半部分的相关叙述。

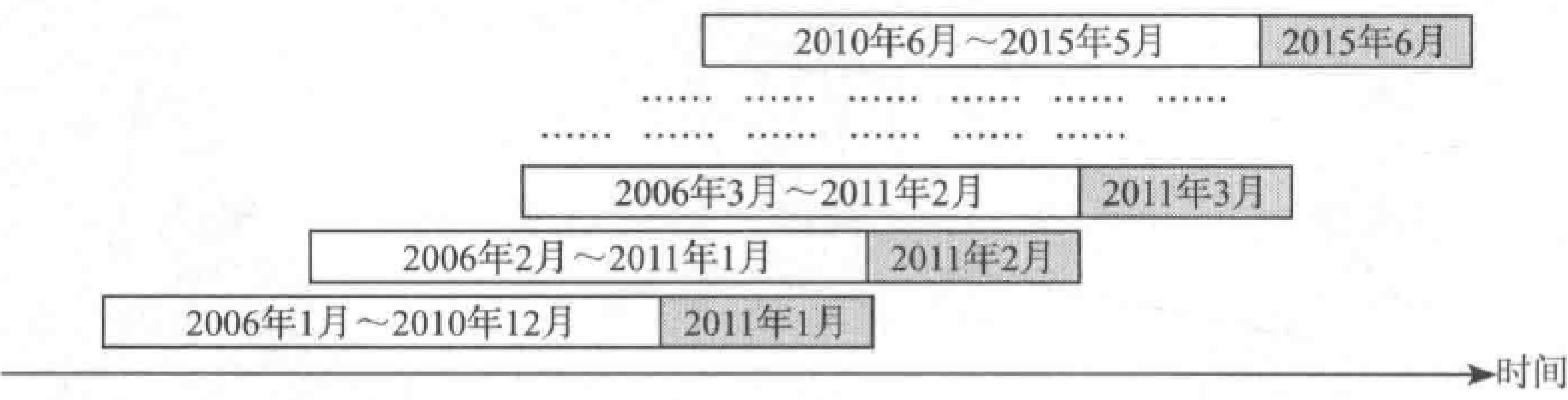


图 8-1 多因子选股策略的推进分析流程

数据方面也和前面所获取的数据一致，同样是自 1991 年 1 月开始直到 2015 年 6 月底结束的国内股票月度数据。所选股票的范围涵盖在深圳证券交易所进行交易的所有 A 股，包括主板、中小企业板和创业板三个板块，以及在上海证券交易所进行交易的所有 A 股。月度数据则包括月末收盘时股票的复权后月度收益率，以及股票的月度流通市值。在进行具体的推进分析之前，为了更好地反映出月度收益率和流通市值之间的线性关系，需要对流通市值进行对数化处理。图 8-1 展示了多因子选股策略的推进分析在这段时间的样本下是如何进行的。1991 年 1 月到 2005 年 12 月的月度数据作为多因子回归模型估计的备用数据，但是实际需要使用的也只有 2005 年 12 月这一个月的数据。这一部分在图中没有具体显示，特此说明。

在推进分析的第一行操作中，白色框为优化期，从 2006 年 1 月到 2010 年

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

12 月一共 60 个月。根据这 60 个月的月度数据，结合 2005 年 12 月的备用数据，就可以得到 60 组用来回归的数据，从而进行多因子回归模型的估计，具体的估计过程可以参考第 5 章 5.4 节中的说明，这里不再赘述。使用白色框的数据完成回归式的估计之后，则将回归式用到灰色框的样本数据上进行交易的模拟。具体来说，就是在 2010 年 12 月收盘的选股时间点上，基于 2010 年 12 月所有股票的月度流通市值和月度收益率数据，使用白色框中得到的回归式计算得到所有股票在 2011 年 1 月的收益率的预测值；然后将股票按照收益率预测值从高到低进行排序，再按顺序等分为若干个部分，每一等份中的股票均按照等权重进行组合，也就是买入占用的资金量相等的组合。在进行完一行的操作之后，将白色框和灰色框都分别向后推进一个月，重复相同的排序和等分过程，直到最后一行为止。

在得到了所有灰色框中的排序结果之后，需要进行的处理实际上与简单优化下的多因子选股完全一致。首先根据每一个等份下长度为 54 的收益率序列，计算出各个等份的平均月度收益情况，进而判断多因子模型预测下一期收益率的整体趋势性。交易策略方面，则根据排序中的第一个等份选择股票，这个序列的股票组合即是每一个月多因子选股策略所选择的股票组合，而该等份的收益率序列也就是策略在每个月下的交易收益。根据这一长度为 54 的收益率序列，可以计算出多因子选股策略在灰色框模拟交易下的净值走势。

灰色框下的研究设置沿用前面简单优化时所采用的设置，即将所有股票按照收益预期大小排序划分为 3 等份、10 等份、30 等份共 3 种情况来加以研究。

8.1.1 多因子选股 3 等分

图 8-2 展示了将所有股票按照多因子模型的预期收益大小划分为 3 等份，每一等份中的股票又按照等权重进行组合时，3 个等份各自的平均月度收益率情况，具体大小分别为 2.38%、1.51% 和 0.60%。第 1 等份与第 2 等份之间相差 0.87%，第 2 等份与第 3 等份之间相差 0.91%，呈现出一个比较

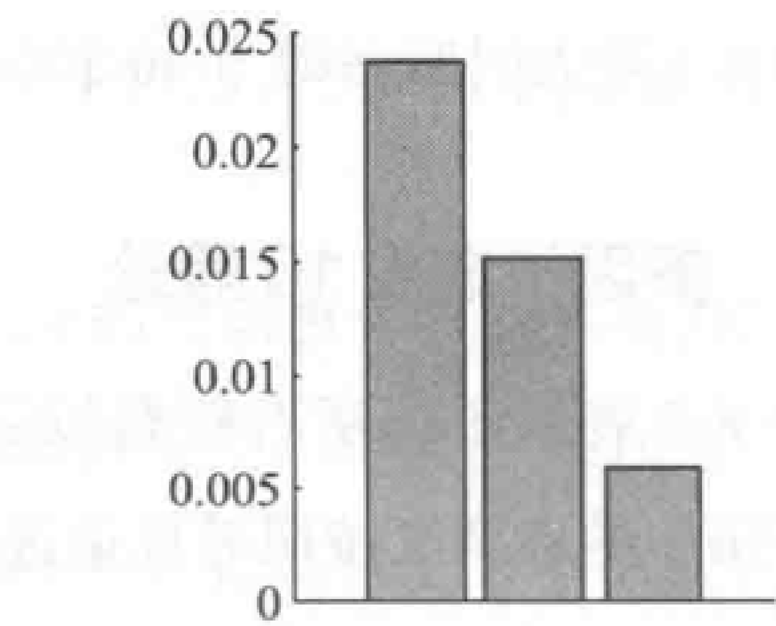


图 8-2 多因子选股 3 等分的收益情况

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

稳定的下降趋势。

确定了逐月的排序和 3 等分之后，就可以得到第 1 等份股票组合的逐月变动情况，也就是多因子选股策略的交易详情。图 8-3 中的实线展示了使用多因子选股策略进行交易的情况下，以 1 为初始值的策略净值走势，虚线则是用来进行对比的同期沪深 300 指数走势。策略净值在 2015 年 6 月收盘时达到 3.5613，四年半的策略收益率为 256.13%，折算下来的年化收益率为 32.61%。

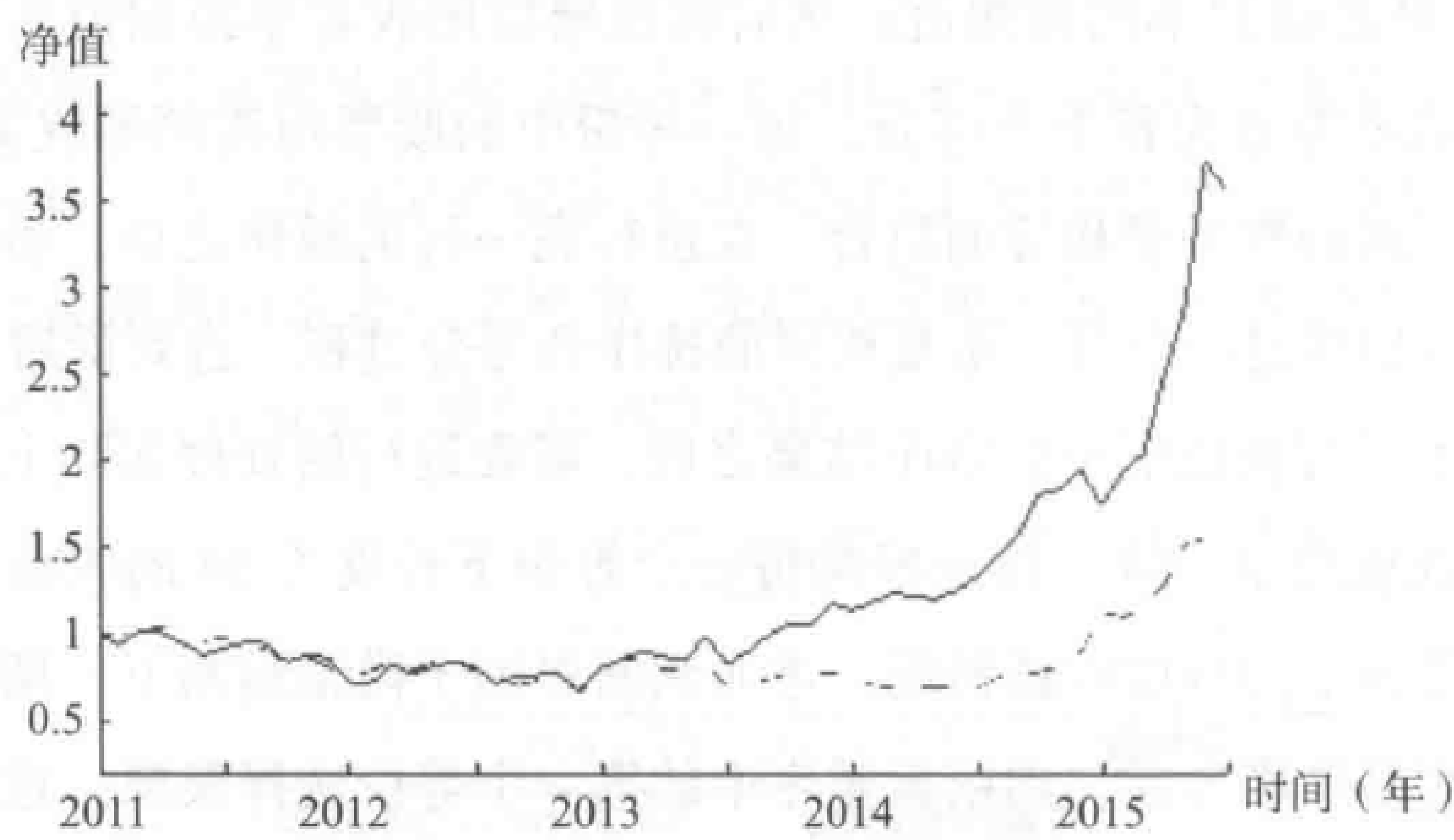


图 8-3 多因子选股前 1/3 的股票组合净值走势

图 8-4 则是逐月地将多因子选股排序前 1/3 的股票组合收益率减去沪深 300 指数的收益率后，所得到的收益率序列形成的净值走势。净值从 2010 年 12 月收盘的 1 开始，在 2015 年 6 月底达到 2.2777，四年半的收益率为 127.77%，折算下来的年化收益率为 20.07%。四年半之间的最大回撤率为 40.54%，回撤的开始点为 2014 年 9 月收盘的 2.3527，结束点则是 2014 年 12 月收盘的 1.3990。同期对比来看，多因子选股简单优化的 3 等分策略下，四年半的收益率为 133.64%，最大回撤率则为 40.50%，且回撤期的时间一致。

8.1.2 多因子选股 10 等分

图 8-5 展示了将所有股票按照多因子模型的预期收益大小划分为 10 等份，每一等份中的股票又按照等权重进行组合时，10 个等份各自的平均月度收益情况。第 1 等份的月均收益率是 10 个等份中的最高值，为 2.71%，第 10 等份的月均收

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

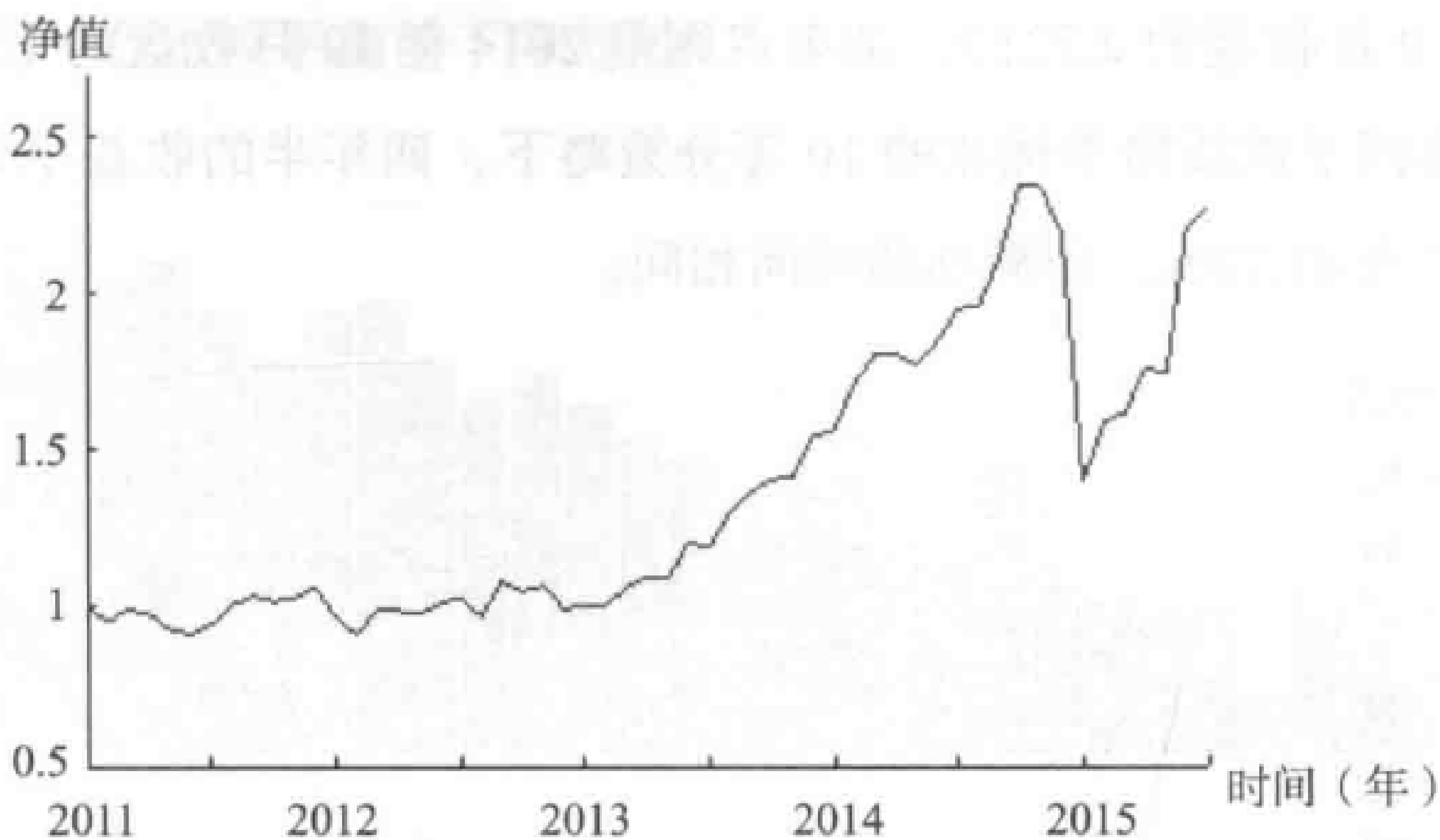


图 8-4 多因子选股前 1/3 的股票组合与沪深 300 收益之差的净值走势

益率则是 10 个等份中的最低值，为 0.29%。其他等份也按照排序关系存在严格的下降趋势，而且下降的幅度比较稳定。这与第 5 章 5.4 节中相关结果的表现是一致的，说明与两个单因子相比，多因子模型可以使等分下的未来收益率具有更平稳的趋势。而由于推进分析本身的优点，这里的结果无疑更加具有说服力。

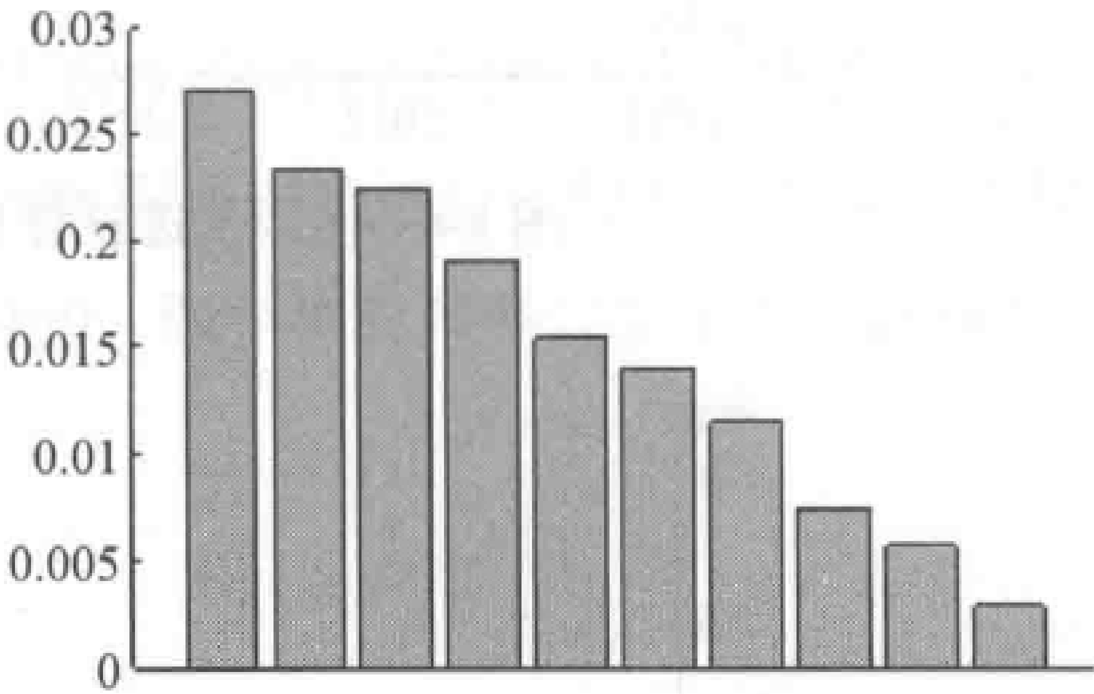


图 8-5 多因子选股 10 等分的收益情况

在确定了逐月的排序和 10 等分之后，也就得到了通过多因子模型选出的第 1 等份股票组合的逐月变动详情。图 8-6 中的实线展示了使用多因子选股策略进行交易时，以 1 为初始值的净值走势情况，虚线则是用来进行对比的沪深 300 指数走势。策略净值在 2015 年 6 月底达到 4.2323，四年半的策略收益率为 323.23%，折算下来的年化收益率为 37.80%。

图 8-7 则是逐月地将多因子选股排序前 1/10 的股票组合收益率减去沪深 300 指数的收益率后，所得到的收益率序列形成的净值走势。净值从 2010 年 12 月收盘的 1 开始，在 2015 年 6 月底达到 2.6757，四年半的收益率为 167.57%，折算下来的年化收益率为 24.45%。四年半之间的最大回撤率为 41.80%，回撤的开始

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

点为 2014 年 9 月收盘的 2.7217，结束点则是 2014 年 12 月收盘的 1.5841。同期对比来看，多因子选股简单优化的 10 等分策略下，四年半的收益率为 177.47%，最大回撤率则为 41.79%，回撤期的时间相同。

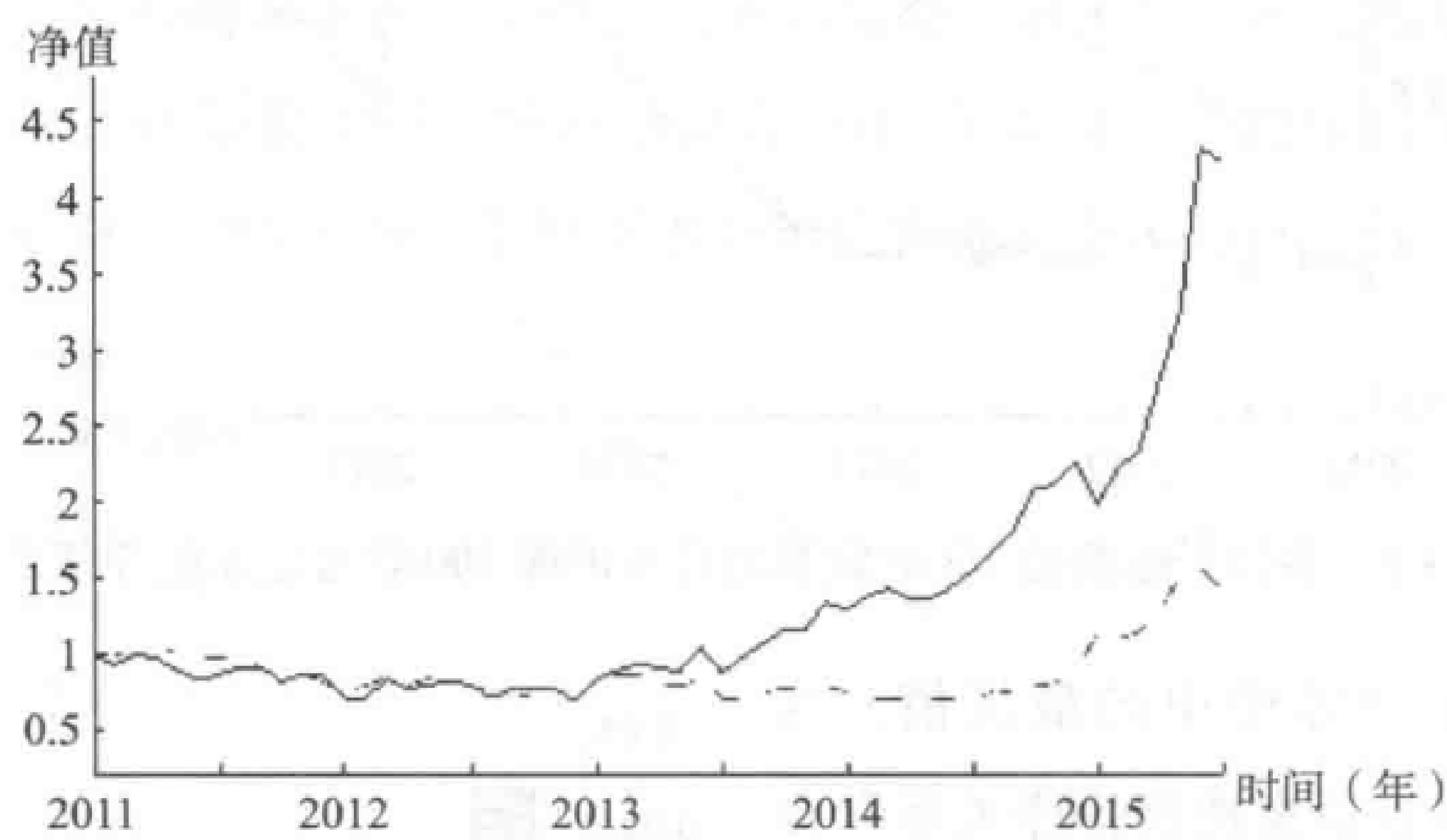


图 8-6 多因子选股前 1/10 的股票组合净值走势

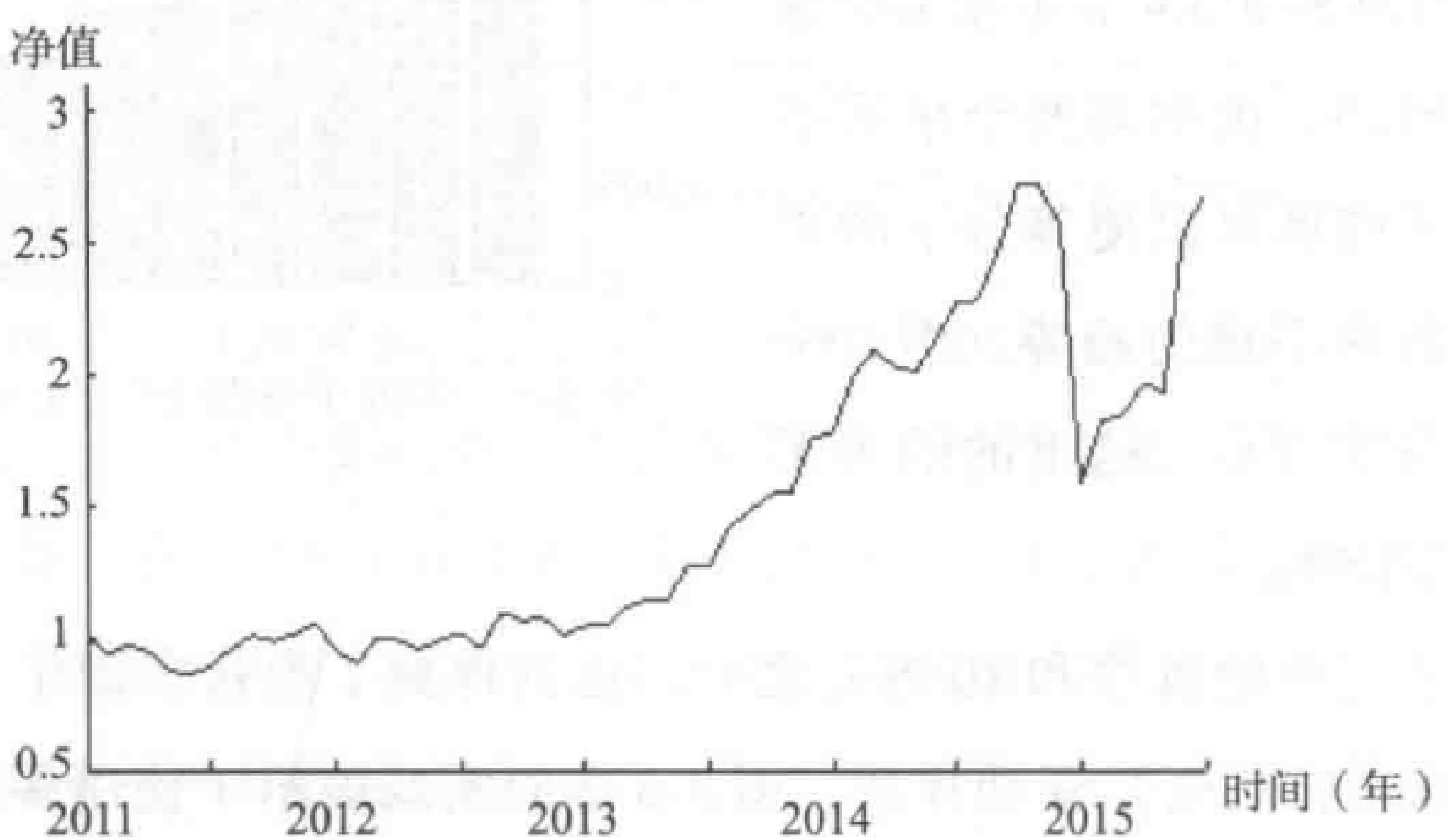


图 8-7 多因子选股前 1/10 的股票组合与沪深 300 收益之差的净值走势

8.1.3 多因子选股 30 等分

图 8-8 展示了将所有股票按照多因子模型的预期收益大小划分为 30 等份，每一等份中的股票又按照等权重进行组合时，30 个等份各自的平均月度收益情况。第 1 等份的月均收益率同样也是 30 个等份中的最高值，为 2.76%，第 30 等份的月均收益率也是 30 个等份中的最低值，为 -0.083%。其他等份的收益情况

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

也基本显示出一个比较平稳的下降趋势。

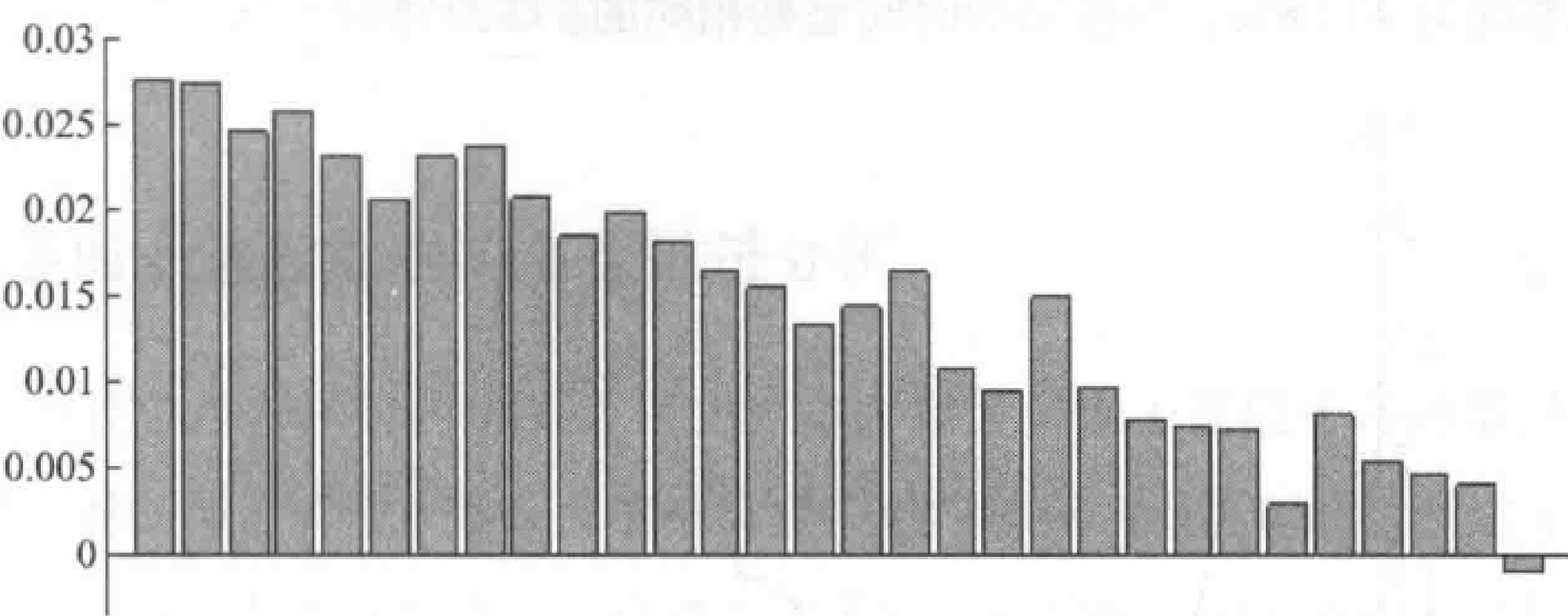


图 8-8 多因子选股 30 等分的收益情况

在确定了逐月的排序和 30 等分之后，也就得到了通过多因子模型选出的第 1 等份股票组合的逐月变动详情。图 8-9 中的实线展示了 30 等分的情况下，以 1 为初始值的多因子选股策略净值走势，虚线则是用来进行比较的沪深 300 指数走势。策略净值在 2015 年 6 月底达到 4.3460，四年半的策略收益率为 334.60%，折算下来的年化收益率为 38.61%。

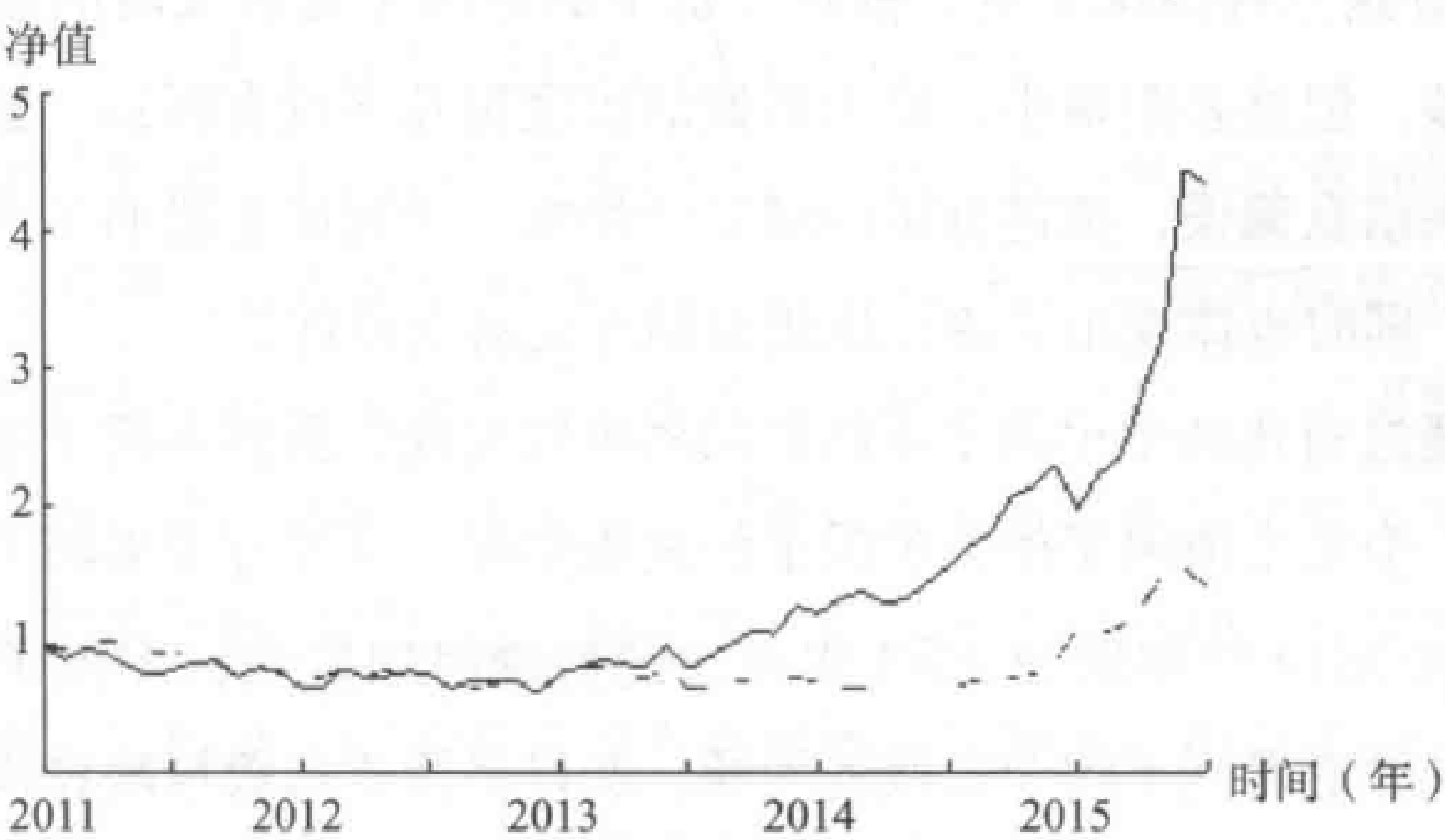


图 8-9 多因子选股前 1/30 的股票组合净值走势

图 8-10 则是逐月地将多因子选股排序前 1/30 的股票组合收益率减去沪深 300 指数的收益率后，所得到的收益率序列形成的净值走势。净值从 2010 年 12 月收盘的 1 开始，在 2015 年 6 月底达到 2.7434，四年半的收益率为 174.34%，折算下来的年化收益率为 25.14%。四年半的最大回撤率为 43.02%，回撤的开始点为 2014 年 10 月收盘的 2.7749，结束点则是 2014 年 12 月收盘的 1.5812。同期对比

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

来看，多因子选股简单优化的 30 等分策略下，四年半的收益率为 191.79%，最大回撤率则为 43.08%，回撤期的时间也是相同的。

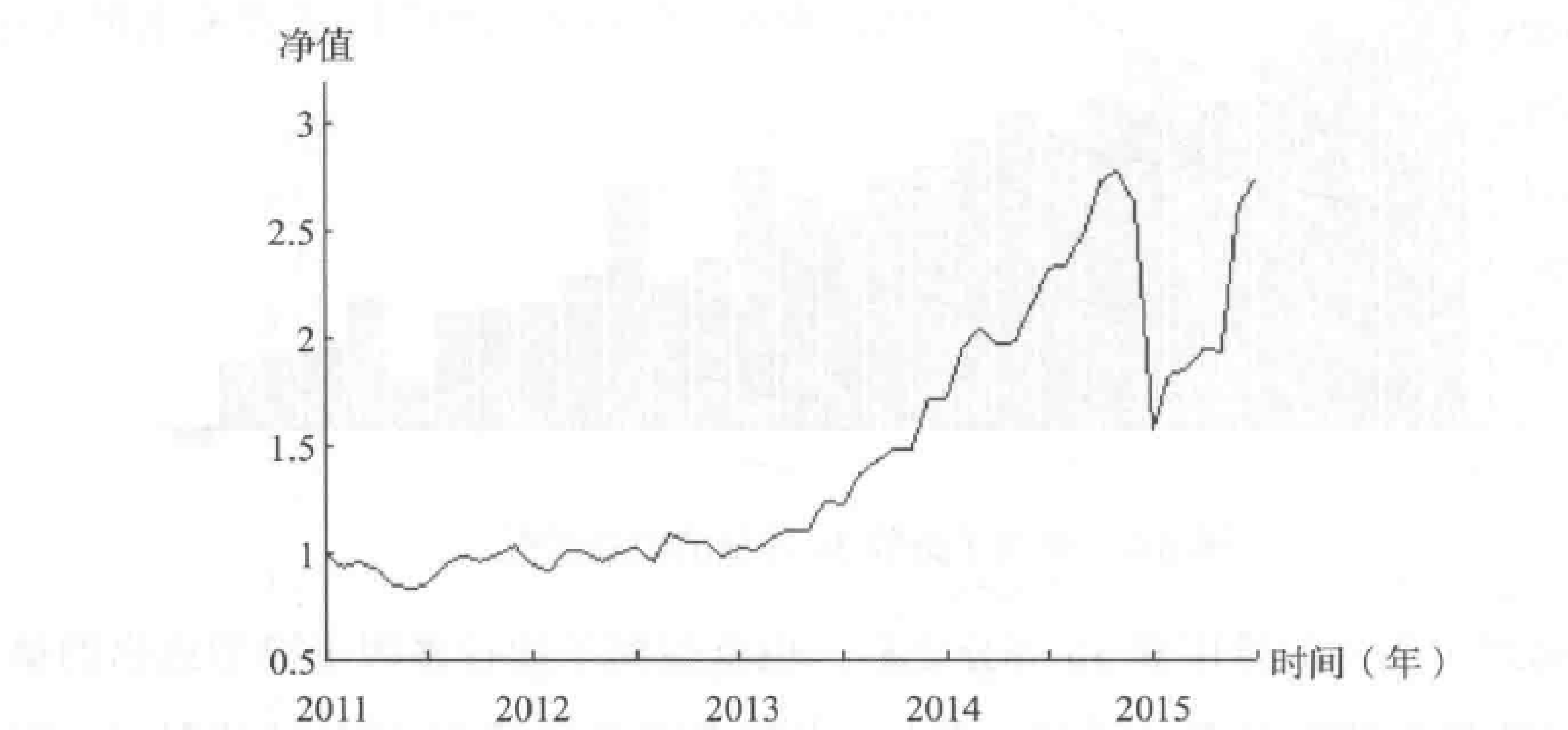


图 8-10　多因子选股前 1/30 的股票组合与沪深 300 收益之差的净值走势

由于本节中多因子选股策略推进分析的模拟交易时间与第 5 章 5.4 节简单优化下的模拟交易时间完全重合，因此可以直接进行相应的对比。综合 3 等分、10 等分、30 等分这三种情况来看，推进分析下的多因子选股策略的盈利表现较之简单优化略差，但是差距很小，最大回撤的程度则几乎没有区别。鉴于简单优化中存在的未来信息偏误，推进分析的结果已经在一定程度上说明了多因子选股策略的有效性，同时也体现出了逐行优化更贴近交易点的优势。

此外，通过对比两个单因子策略的结果可以发现，虽然多因子选股策略中包含市值因子，但是市值因子作为单因子的表现实际上是要好于多因子选股的。这主要是由于在 2014 年年底和 2015 年年初的特殊时间段内，多因子选股策略中的两个因子同时遭遇到了因子失效的问题，所以多因子选股策略在没有降低最大回撤风险的情况下，收益情况却被两个因子平均化了，进而形成了这种退化的结果。但是在真实的策略研发过程中，即使在样本中出现了这样的问题，作者同样建议在多因子选股策略中尽可能多地吸收有效因子。只要研究人员通过逻辑判断出多个因子之间没有非常强的内在相关性，例如本书例子中的市值因子和上个月收益反转因子，就应当把多个有效因子的结合视为一个可能更优的选择来看待。这样的做法源于多因子选股策略本身的内在经济直觉，而不是数据实证的结果。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

虽然本书整个篇幅都在讲解数据的实证研究，但是在作者的观念里，经济直觉的重要性其实是要高于纯粹的数据研究结果的。

8.2 多因子选股策略的多层推进分析

上一节中作者以国内所有 A 股为样本介绍了多因子选股策略的推进分析流程，并发现在特定的样本期以及策略设置下可以得到一个正面的盈利结果。在这部分的研究中，由于回归模型来自选股因子的交易逻辑，因此并没有具体的设置参数需要提前确定。但是正如第 6 章中提到过的，训练样本的大小也就是图 8-1 中白色框的长度，作为推进分析本身的设置参数是需要在分析之前人为给定的。在这一节中，作者就主要针对训练样本大小这一参数设置进行多因子选股策略的双层推进分析，从而进一步判断多因子选股策略的实际盈利能力。当然，这也同时为读者提供了一个可以参考的多层推进分析示例，特别是针对推进分析自身设置所进行的多层推进分析示例。至于多因子选股策略的构造形式，读者可以参考上一节以及第 5 章 5.4 节的相关内容，这里不再赘述。

多层推进的数据也和前面所获取的数据一致，样本涵盖在深圳证券交易所和上海证券交易所进行交易的所有 A 股，具体包括月末收盘时股票的复权后月度收益率，以及股票的月度流通市值。在进行策略研究之前，为了更好地反映出月度收益率和流通市值之间的线性关系，需要对流通市值进行对数化处理。图 8-11 描述了多因子选股策略双层推进分析中第二层的数据划分和运行情况，而为了简化示意图，推进分析第一层的设置没有在图中进行展示。但是值得注意的是，这一节研究中第一层推进分析的具体情况与第 7 章 7.5 节中的情况又有所不同。在自回归择时策略的双层推进分析中，第二层推进分析主要针对的是自回归模型的阶数，因此第一层白色框的长度依然保持为 1 年没有改变，第一层推进分析的具体形式也相对比较固定。但是在这里，第一层白色框的长度正是第二层推进分析所针对的检验对象，因此图 8-11 中没有展示的第一层推进分析的结构其实是随着第二层的检验过程而不断变动的。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net



图 8-11　多因子选股策略的双层推进分析流程

下面进行具体的说明。在第二层推进分析的第一行操作中，浅灰色框为第一层白色框长度设置的优化期，同时也是第一层推进分析的模拟交易期，从 2008 年 1 月开始，直到 2010 年 12 月结束，共计 36 个月。对白色框长度的优化采用完全的网格搜索，也就是将长度分别设置为 1 ～ 120，使用与上一节中雷同的推进分析流程逐个测试不同长度下的模拟交易情况，进而选取盈利能力最强的白色框长度设置作为最优结果。因此，第一行操作中没有具体显示的白色框最远回溯到 120 个月之前，也就是 1997 年 1 月。再加上估计多因子回归模型所需要用到的前一个月的备用数据，全部的双层推进分析所使用的数据实际从 1996 年 12 月开始。需要说明的是，由于中国股票市场的迅速发展，较早的样本数据与当前的股票数据在内在特征上是否具有较大差异，显然是一个值得考虑的问题。因此在本节这种样本期较长的情况下，有必要对研究结果持有更高的怀疑态度。

在确定了第一行的最优白色框长度之后，就可以将该长度设置固定到深灰色框的多因子回归模型之上，进行相似的第一层推进分析操作，也就是根据前期数据估计出回归模型系数之后，在 2010 年 12 月收盘的选股时间点上，基于 2010 年 12 月所有股票的月度流通市值和月度收益率数据，使用得到的回归式计算出所有股票在 2011 年 1 月的收益率的预测值。然后将股票按照收益率预测值从高到低进行排序，再按顺序等分为若干个部分，每一等份中的股票又按照等权重进行组合，也就是买入占用的资金量相等的组合。在进行完一行的操作之后，将所有的数据框都分别向后推进一个月，重复相同的研究过程，直到最后一行为止。

在得到了所有深灰色框中的排序结果之后，需要进行的处理实际上也就与其他因子选股研究流程完全一致了。首先根据每一个等份下长度为 54 的收益率序列，计算出各个等份的平均月度收益情况，进而判断多因子模型预测下一期收

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

益率的整体趋势性。交易策略方面，则根据排序中的第一个等份选择股票，这个序列的股票组合即是每一个月多因子选股策略所选择的股票组合，而该等份的收益率序列也就是策略在每个月下的交易收益。根据这一长度为 54 的收益率序列，可以计算出多因子选股策略在深灰色框模拟交易下的净值走势。

深灰色框下的研究设置依然沿用前面研究选股策略时所采用的设置，即将所有股票按照收益预期大小排序划分为 3 等份、10 等份、30 等份共三种情况来加以研究。

8.2.1 多因子选股 3 等分

图 8-12 展示了将所有股票按照多因子模型的预期收益大小划分为 3 等份，每一等份中的股票又按照等权重进行组合时，3 个等份各自的平均月度收益率情况。具体的收益率大小分别为 2.25%、1.49%、0.75%。第 1 等份与第 2 等份之间相差 0.76%，第 2 等份与第 3 等份之间相差 0.74%，呈现出了一个比较稳定的下降趋势。

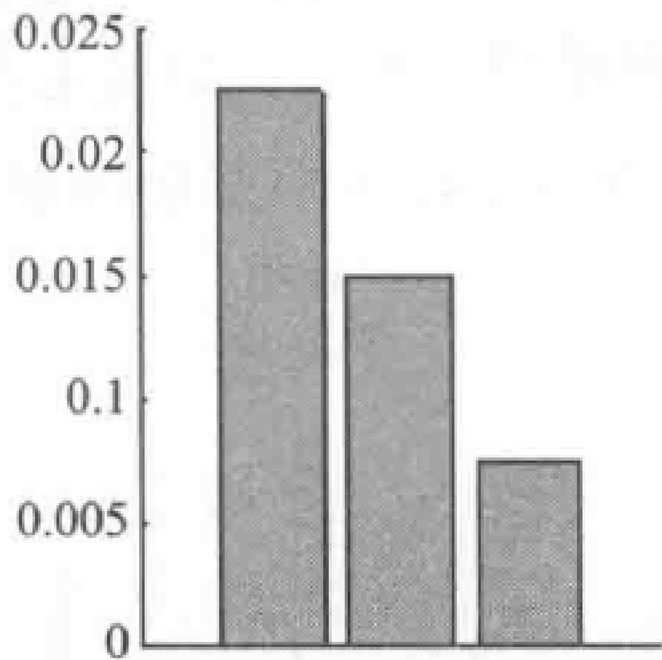


图 8-12 多因子选股 3 等分的收益情况

确定了逐月的排序和 3 等分之后，就可以得到第 1 等份股票组合的逐月变动情况，也就是多因子选股策略的交易详情。图 8-13 中的实线展示了使用多因子选股策略进行交易的情况下，以 1 为初始值的策略净值走势，虚线则是用来进行对比的同期沪深 300 指数走势。策略净值在 2015 年 6 月收盘时达到 3.3271，四年半的策略收益率为 232.71%，折算下来的年化收益率为 30.62%。

图 8-14 则是逐月地将多因子选股排序前 1/3 的股票组合收益率减去沪深 300 指数的收益率后，所得到的收益率序列形成的净值走势。净值从 2010 年 12 月收盘的 1 开始，在 2015 年 6 月底达到 2.1358，四年半的收益率为 113.58%，折算下来的年化收益率为 18.37%。四年半之间的最大回撤率为 40.62%，回撤的开始点为 2014 年 9 月收盘的 2.1933，结束点则是 2014 年 12 月收盘的 1.3024。同期对比来看，上一节推进分析的 3 等分策略下，四年半的收益率为 127.77%，最大

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

回撤率则为 40.54%，回撤期的时间是一样的。

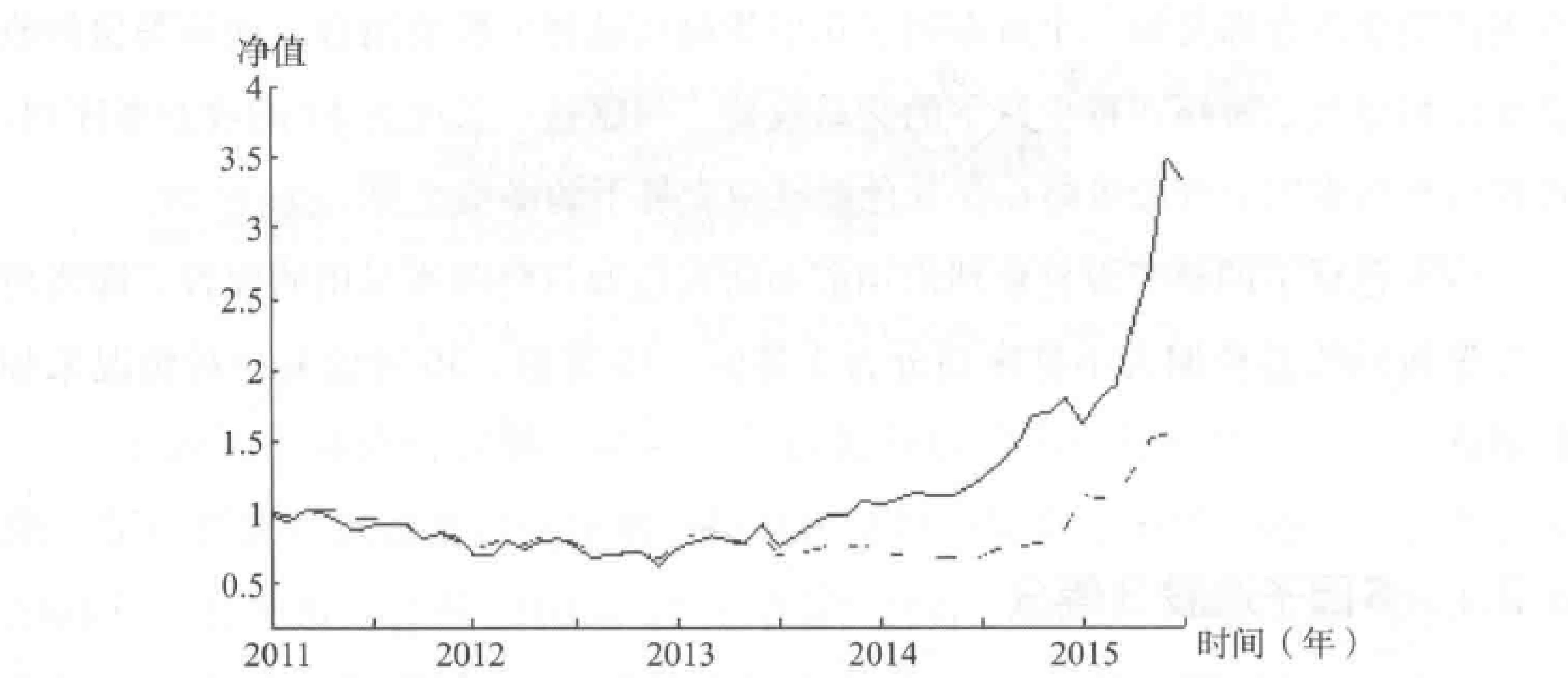


图 8-13 多因子选股前 1/3 的股票组合净值走势

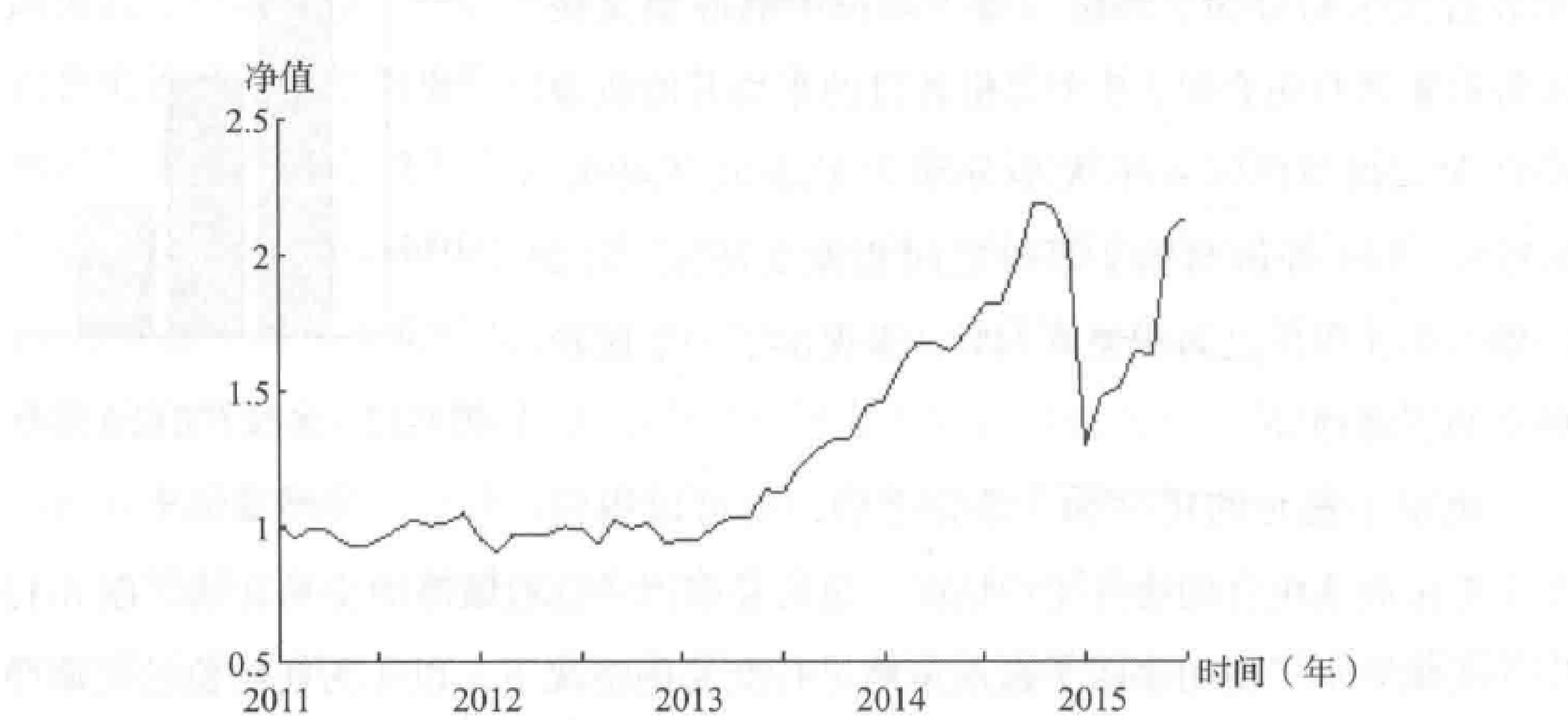


图 8-14 多因子选股前 1/3 的股票组合与沪深 300 收益之差的净值走势

8.2.2 多因子选股 10 等分

图 8-15 展示了将所有股票按照多因子模型的预期收益大小划分为 10 等份，每一等份中的股票又按照等权重进行组合时，10 个等份各自的平均月度收益情况。第 1 等份的月均收益率是 10 个等份中的最高值，为 2.42%，第 10 等份的月均收益率则是 10 个等份中的最低值，为 0.54%。其他等份也按照排序关系存在严格的下降趋势，而且下降的幅度比较稳定。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

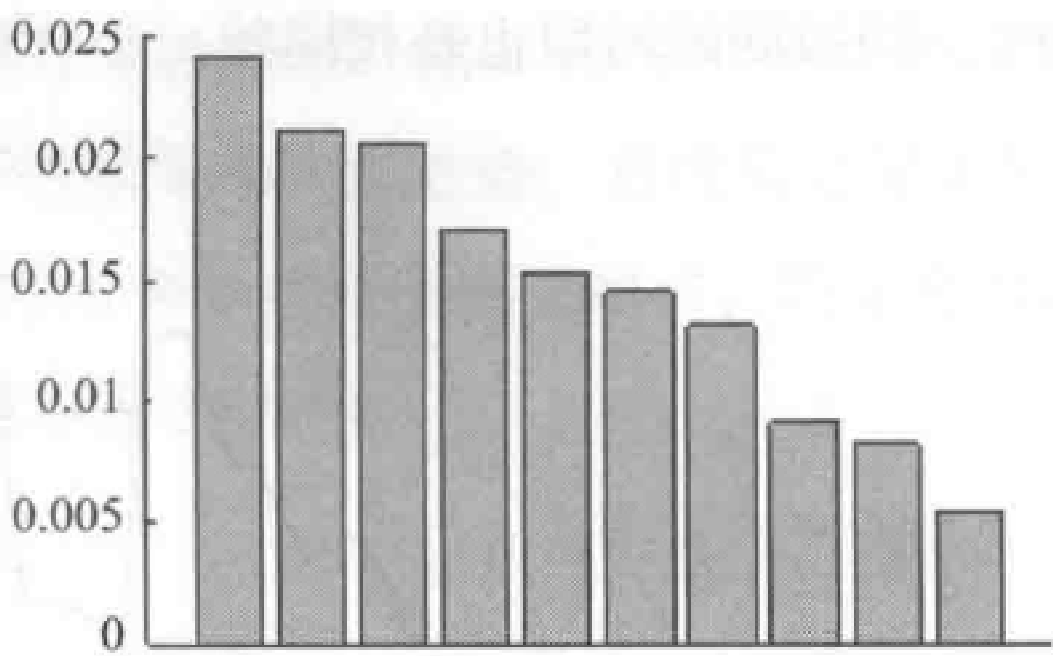


图 8-15　多因子选股 10 等分的收益情况

在确定了逐月的排序和 10 等分之后，也就得到了通过多因子模型选出的第 1 等份股票组合的逐月变动详情。图 8-16 中的实线展示了使用多因子选股策略进行交易时，以 1 为初始值的净值走势情况，虚线则是用来对比的沪深 300 指数走势。策略净值在 2015 年 6 月底达到 3.6319，四年半的策略收益率为 263.19%，折算下来的年化收益率为 33.19%。

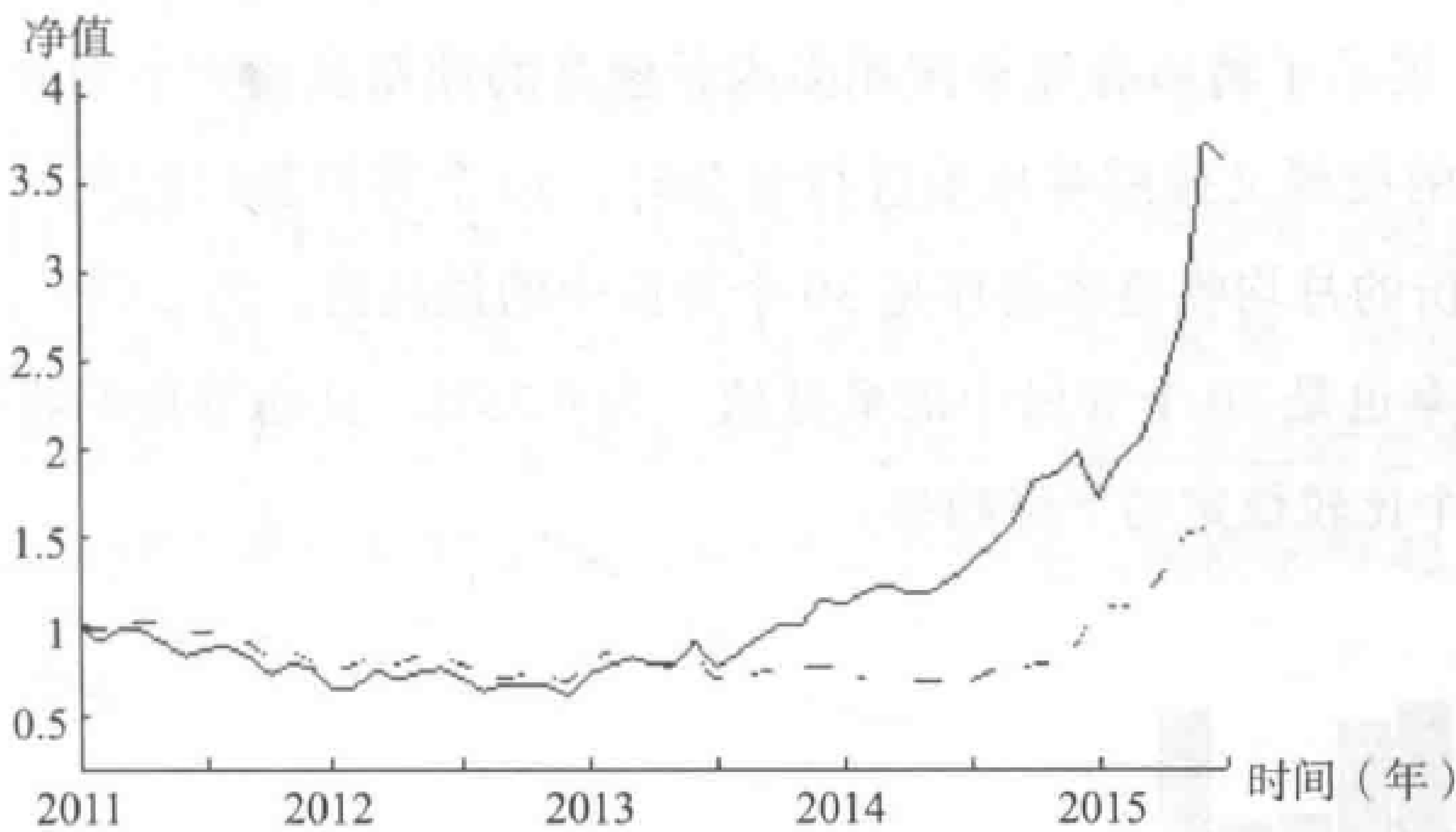


图 8-16　多因子选股前 1/10 的股票组合净值走势

图 8-17 则是逐月地将多因子选股排序前 1/10 的股票组合收益率减去沪深 300 指数的收益率后，所得到的收益率序列形成的净值走势。净值从 2010 年 12 月收盘的 1 开始，在 2015 年 6 月底达到 2.3012，四年半的收益率为 130.12%，折算下来的年化收益率为 20.35%。四年半之间的最大回撤率为 42.29%，回撤的开始点为 2014 年 9 月收盘的 2.4052，结束点则是 2014 年 12 月收盘的 1.3881。同期对比来看，上一节推进分析的 10 等分策略下，四年半的收益率为 167.57%，

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

最大回撤率则为 41.80%，回撤期的时间也是相同的。

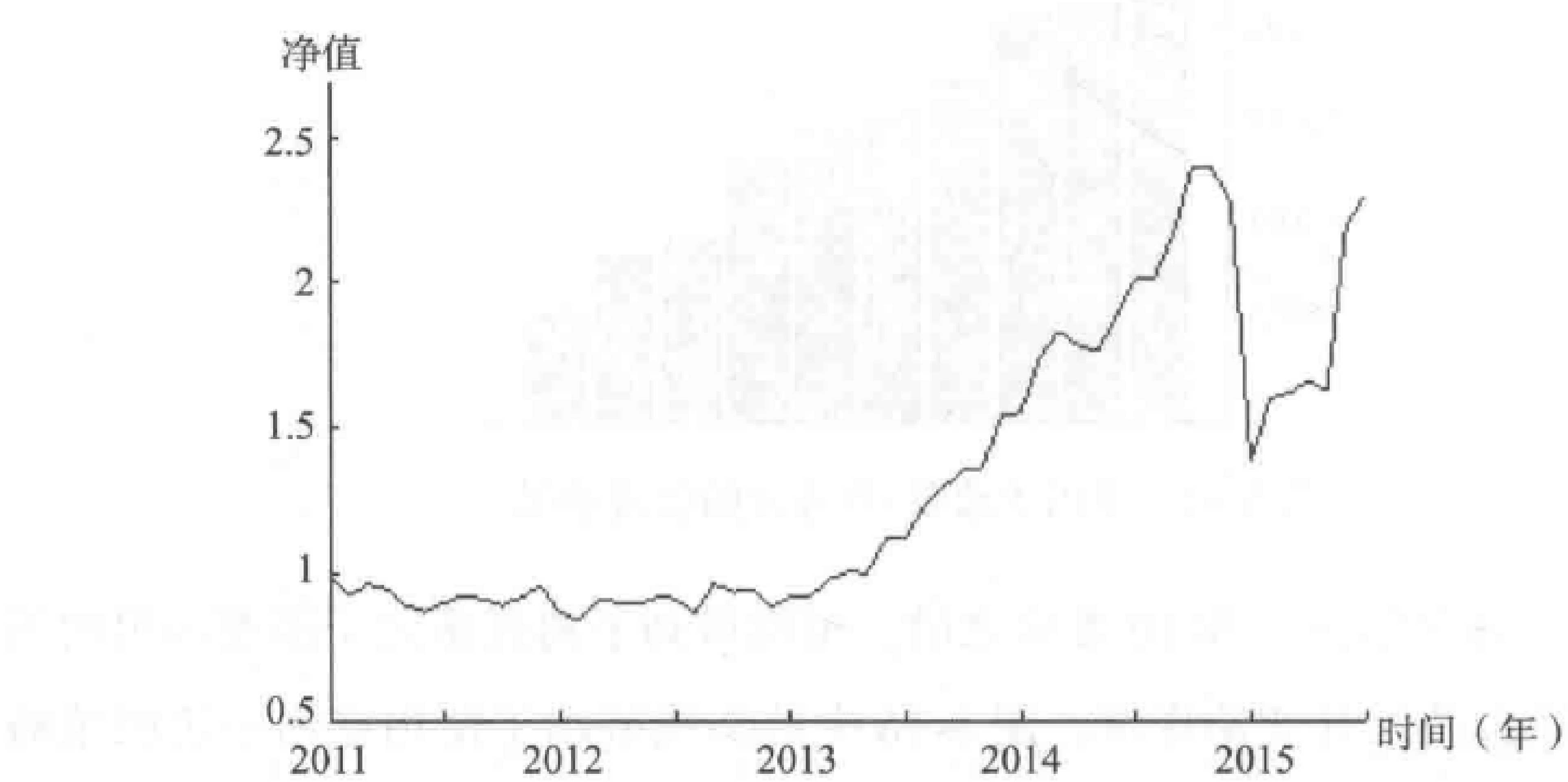


图 8-17　多因子选股前 1/10 的股票组合与沪深 300 收益之差的净值走势

8.2.3　多因子选股 30 等分

图 8-18 展示了将所有股票按照多因子模型的预期收益大小划分为 30 等份，每一等份中的股票又按照等权重进行组合时，30 个等份各自的平均月度收益情况。第 1 等份的月均收益率同样是 30 个等份中的最高值，为 2.59%，第 30 等份的月均收益率也是 30 个等份中的最低值，为 0.25%。其他等份的收益情况也基本显示出一个比较稳定的下降趋势。

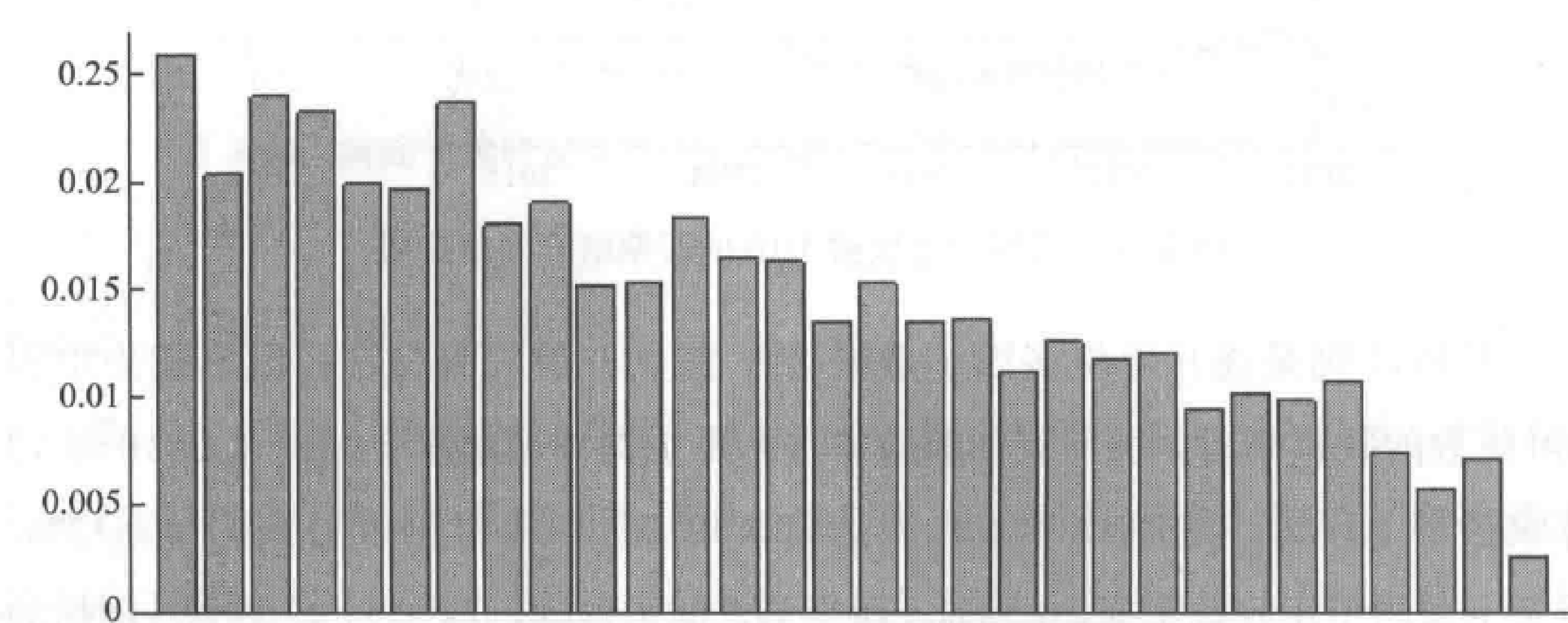


图 8-18　多因子选股 30 等分的收益情况

在确定了逐月的排序和 30 等分之后，也就得到了通过多因子模型选出的第

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

1 等份股票组合的逐月变动详情。图 8-19 中的实线展示了 30 等分的情况下，以 1 为初始值的多因子选股策略净值走势，虚线则是用来进行比较的沪深 300 指数走势。策略净值在 2015 年 6 月底达到 3.9847，四年半的策略收益率为 298.47%，折算下来的年化收益率为 35.96%。

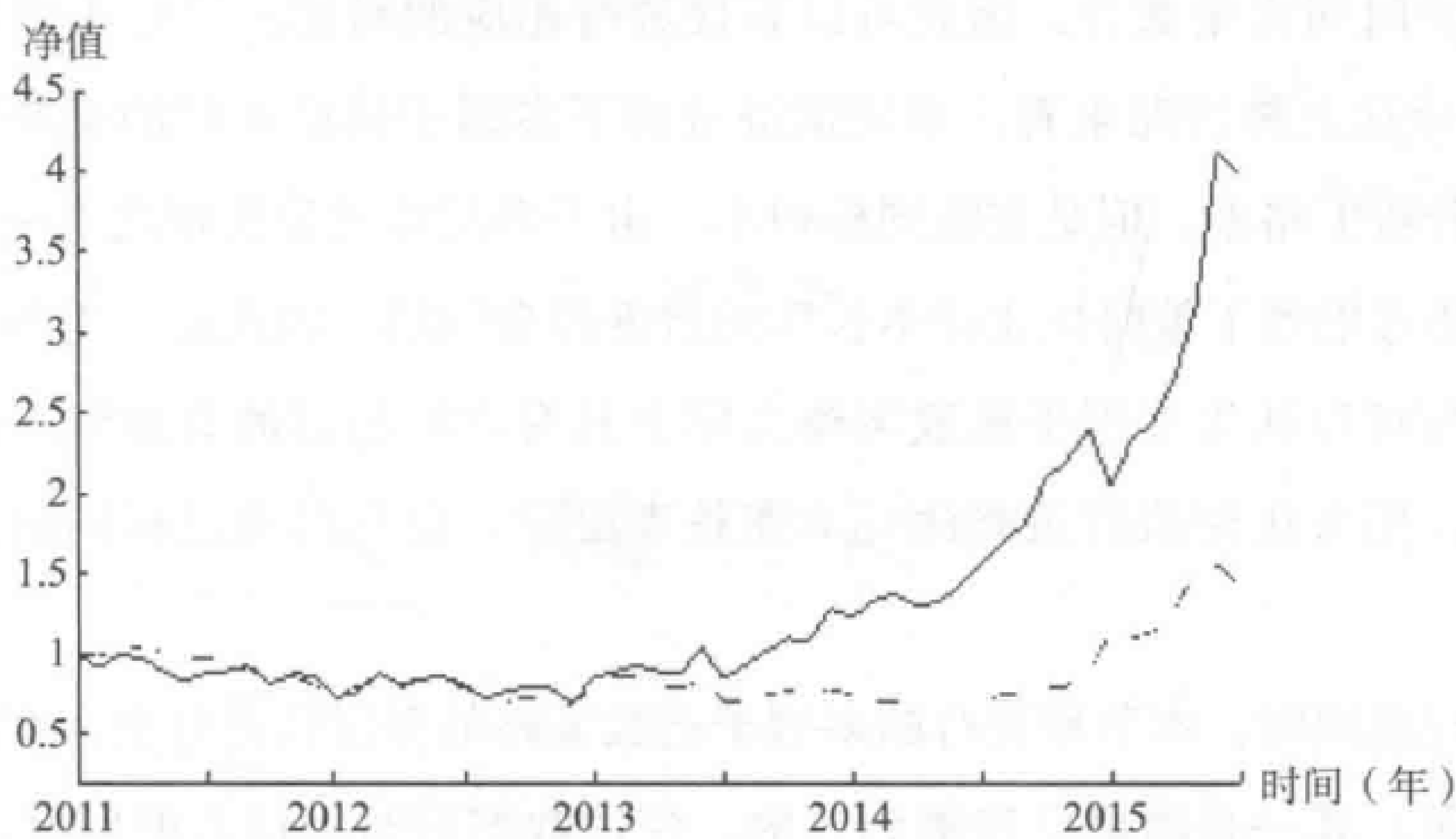


图 8-19 多因子选股前 1/30 的股票组合净值走势

图 8-20 则是逐月地将多因子选股排序前 1/30 的股票组合收益率减去沪深 300 指数的收益率后，所得到的收益率序列形成的净值走势。净值从 2010 年 12 月收盘的 1 开始，在 2015 年 6 月底达到 2.5080，四年半的收益率为 150.80%，折算下来的年化收益率为 22.67%。四年半之间的最大回撤率为 42.51%，回撤的

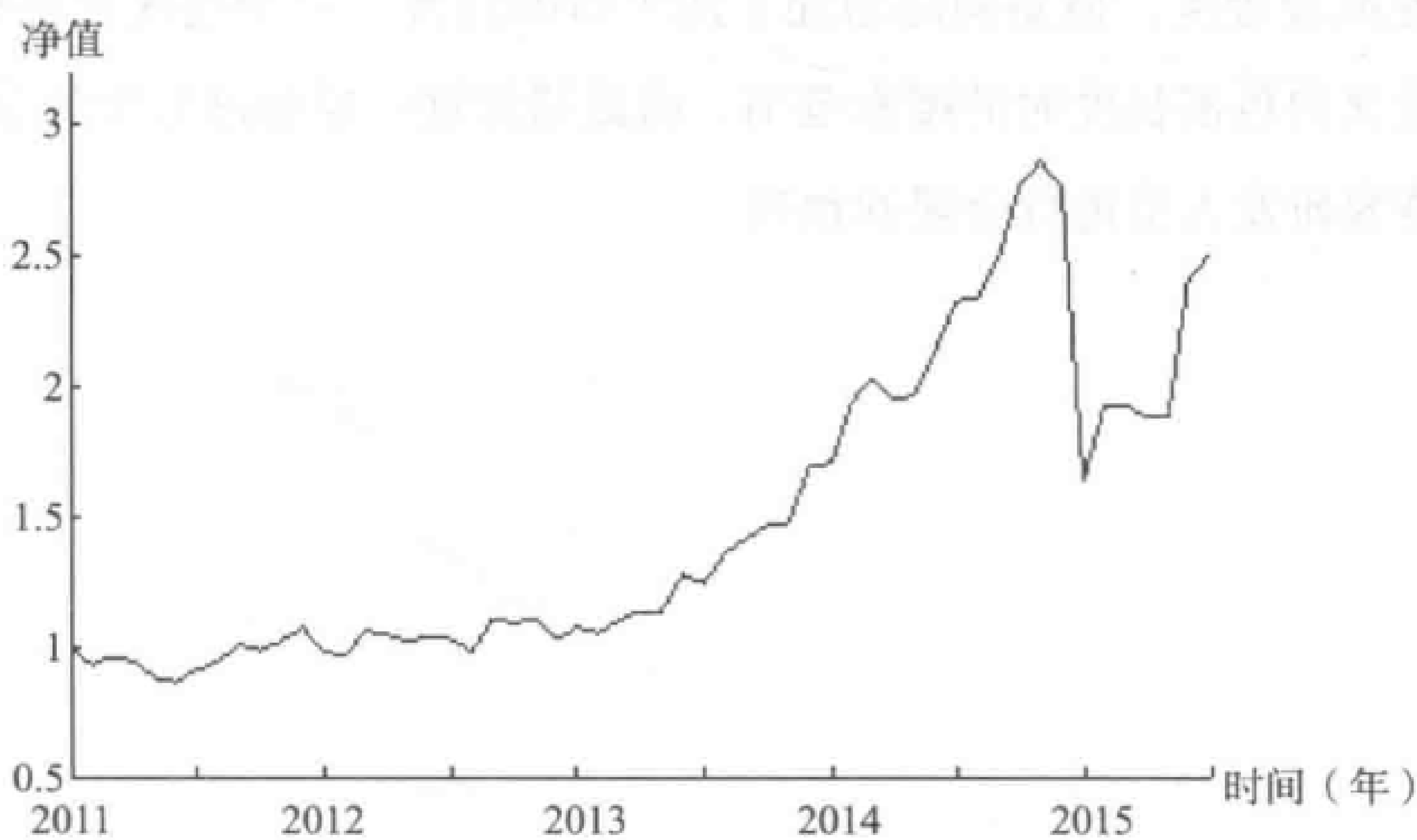


图 8-20 多因子选股前 1/30 的股票组合与沪深 300 收益之差的净值走势

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

开始点为2014年10月收盘的2.8613，结束点则是2014年12月收盘的1.6450。同期对比来看，上一节推进分析的30等分策略下，四年半的收益率为174.34%，最大回撤率则为43.02%，回撤期的具体时间也是相同的。

由于本节中多因子选股策略多层推进分析的模拟交易时间与上一节推进分析的模拟交易时间完全重合，因此可以直接进行相应的对比。综合3等分、10等分、30等分这三种情况来看，多层推进分析下多因子选股策略的交易表现比单纯的推进分析下略差，但是差距同样很小。由于多层推进分析较之上一节的推进分析进一步考虑到了策略优化样本长度的过度拟合问题，因此这一差距是可以接受的，而且可以认为多因子选股策略实际上具有比较稳健的有效性。在实际交易过程中，用来优化执行策略的样本期长度设定，也可以通过相同的方法实时得到。

不过与此同时，本节所进行的多因子选股策略的多层推进分析存在两个需要注意的问题。其一是前面已经提到过的，使用数据的时间过长的问题，从1996年12月收盘开始的前面几年内，中国股市实际上还在不断地成熟与改进中，这一部分数据的应用会在多大程度上影响整个回溯测试过程的内在一致性，是需要研发人员进行主观判断的问题。第二个问题在第6章介绍多层推进分析时也曾进行过阐述，就是这种方法会在检验一些参数设置的过度拟合的同时，引入新的参数设置。在这一节的具体案例中，第二层推进分析可以用来研判白色框长度这一设定的过度拟合程度，但是同时增加了两个新的设置，一个是浅灰色框的长度，另一个是优化白色框长度时的搜索范围。因此增加这一层推进分析的意义到底有多大，仍需要研发人员进行合理的预判。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

第9章

风　险

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

9.1 常用的风险度量

在第 2 章对量化交易策略的研发流程进行介绍时，作者说明了收益和风险这两个特征对于策略构建的重要作用与意义。第 4 ～ 8 章则给出了若干交易策略的实例，但是这些策略的构建基本上仅由收益特征所推动。在这一章中作者将在量化交易策略的背景下阐述与风险相关的内容，从而为后面的章节做出基础性的铺垫。

不同于收益，风险实际上是一个较为抽象的概念，学术界和业界对它都缺乏一个约定俗成、统一的定义，而对于单个的个人而言，对风险的理解更可能千差万别。好在存在一个笼统但是涵盖较广的风险定义，可以包容大部分人对于风险的认识，即风险是事件未来可能结果的不确定性。本书将从这一内涵出发，尝试梳理一下量化交易策略研发下的风险特征以及相应的度量方法。

在量化交易策略的研发过程中，我们所关注的事件是运行策略所形成的净值走势情况，而更进一步分解时间序列上的净值走势，实际上我们关注的结果正是策略所带来的收益。在这样的背景下，收益的不确定性是可以用来描述交易策略所固有的风险特征的。

从前面的内容中可以发现，对收益本身的追求与优化可以形成量化交易策略，收益在这个层面上是有益的、正面的。但是收益的不确定性则相反，当这种不确定性的程度增加时，量化交易策略的最终结果会呈现出不稳定的特征，这种不稳定性也就是风险所带来的缺陷。同时，由于量化交易策略中优化行为的存在，因此这种不确定性会导致优化等数学处理的失效，也就是第 3 章所提到过的过度拟合等问题，这也是风险的负面效应在量化交易策略下的一个实际体现。

在研发量化交易策略时，由于要使用各种数量化手段来进行模型构建和策略描述，因此选定一个精准的风险度量方式是必须进行的工作。而度量收益不确定性的方式多种多样，因此在实际的研发过程中，往往要根据实际情况进行合理的选择。

对收益不确定性描述最为清晰的应该是收益率的概率分布，其在数学定义上包含了可能的结果以及概率，虽然较为复杂但是细节丰富，用来呈现收益的不

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

确定性十分合适。图 9-1 是一个实际的例子，展示了推进分析下自回归策略也就是第 7 章 7.4 节中的交易策略的日收益率分布情况，其中柱状图反映了日收益率的经验分布。观察图 9-1，可以了解到策略收益率的大部分可能情况以及相应的概率。

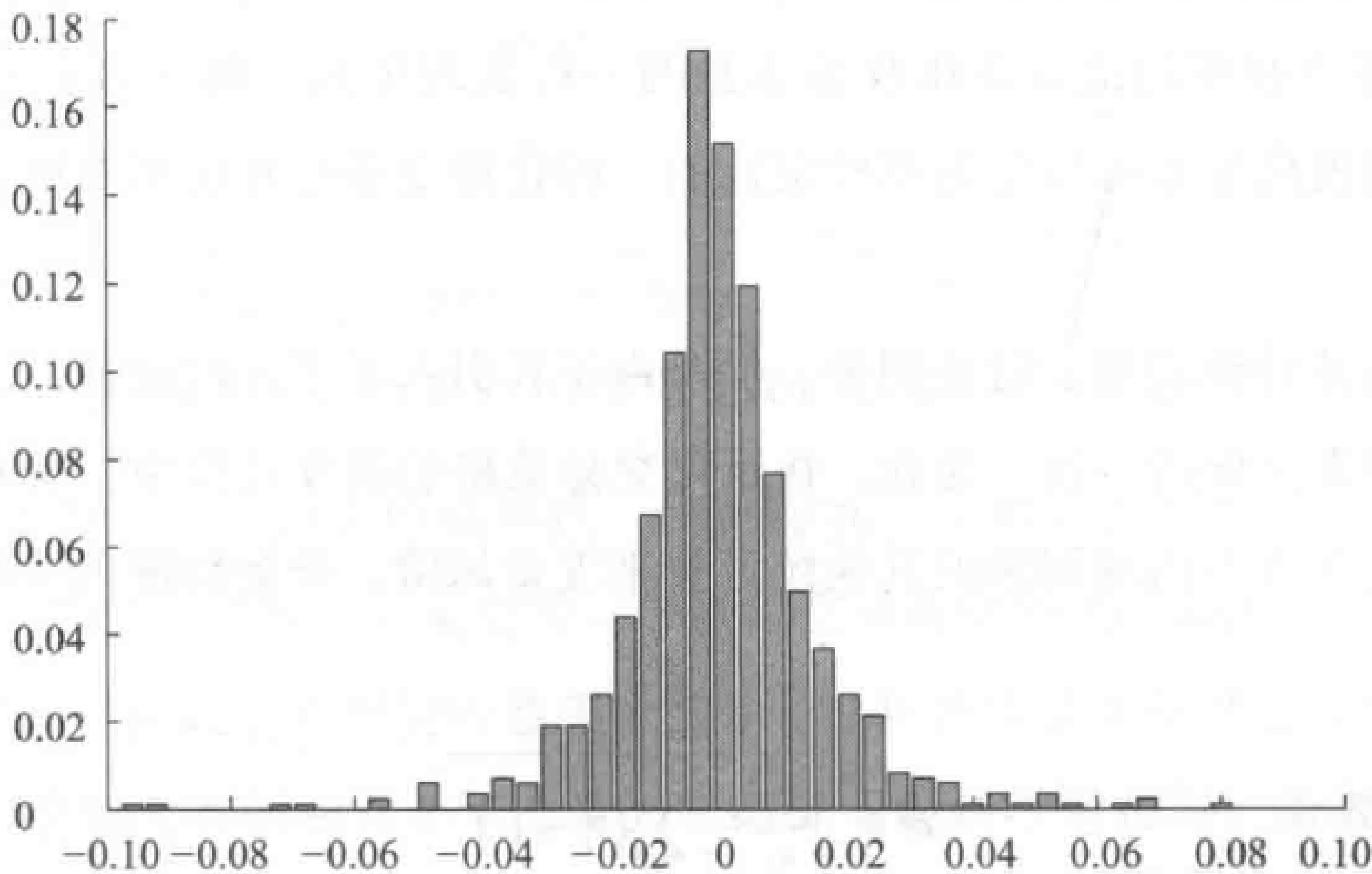


图 9-1　推进分析下自回归策略的收益率分布

然而有一个问题需要注意，虽然收益率的分布可以完整地描述收益的不确定性，但是由于其表现过于细致，往往并不能很好地运用于量化交易策略研发的数学处理过程中。例如，当使用收益分布作为风险度量运用于策略目标最优化时，这种过于细节化的处理会使得优化问题的复杂程度大大增加甚至无法实现。

因此，分布情况这种风险度量方式虽然具有自身的优点，但是一般不作为优化函数之类的表现形式，而仅仅出现在策略结果事后的风险分析中。更多时候，量化交易策略的整个研发过程需要使用一种简洁但有效的风险度量，从而使得风险度量能够同时能够满足实现风险特征表现和事后风险描述两种使用形式的要求。

方差正是这样的一种风险度量方式，其刻画了随机变量的取值对于其数学期望的离散程度。实际上，方差本是一个概率统计上的概念，马科维茨在 1952 年的开创性论文中使用了这一统计量作为投资组合风险的度量形式，其在金融研究中的应用也随之越来越多。仍以推进分析下的自回归策略为例，设推进分析中模

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

拟交易期内第 i 天的策略收益率为 r_i，模拟交易期的长度为 N（该例中 $N=845$），则模拟交易样本下的方差 σ^2 可以计算如下

$$\sigma^2=\frac{1}{N-1}\sum_{i=1}^{N}(r_i-\mu)^2$$

式中，μ 为收益率序列的均值。

如果将收益率均值 μ 看作收益本身的一种表现形式，那么方差这种在稳定收益之上的偏离度就可以作为不确定性的一种合理度量，并用来表现交易策略的风险。

虽然方差计算简便、特性明显，但是由于其引入了平方的设置，因此在量纲上与收益率无法保持一致。为此，在量化交易策略的研发过程中，研究人员往往会使用一些与方差内涵同源的其他统计量来度量风险，例如标准差（σ），其计算方法如下

$$\sigma=\sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(r_i-\mu)^2}$$

实际上就是在样本方差的基础上进行了开平方处理，从而与收益率的量纲保持相同。此外还有如标准误之类的统计量也时有应用，读者可以参阅对此叙述更为详细的概率统计方面的资料。

虽然方差和标准差具有清晰的统计特性，而且在策略模型的构建过程中具有较好的契合度。由于其源于概率统计领域，因此在实际研发量化交易策略的时候，方差和标准差往往对交易员而言缺乏可以直观理解的现实意义。为了达到这一目的，研发人员会参考量化交易策略的最大回撤率这一指标来判断策略的风险。

最大回撤率是一个更偏向于对量化交易策略实际表现进行描述的风险度量形式，具体定义为一段时间内的任一时点向后推进，策略净值走到最低点时的净值回撤幅度的最大值。仍然以第 7 章 7.4 节中的自回归策略交易结果为例，在三年半的模拟交易期内，净值的最大回撤率为 25.92%，回撤的发生从 2014 年 12 月 22 日收盘的 2.2796 开始，一直到 2015 年 2 月 25 日收盘时的 1.6887 结束，图 9-2 基于自回归策略的净值走势图展示了其最大回撤的具体情况，其中的两条横

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

线分别标明了最大回撤的起点和终点位置。

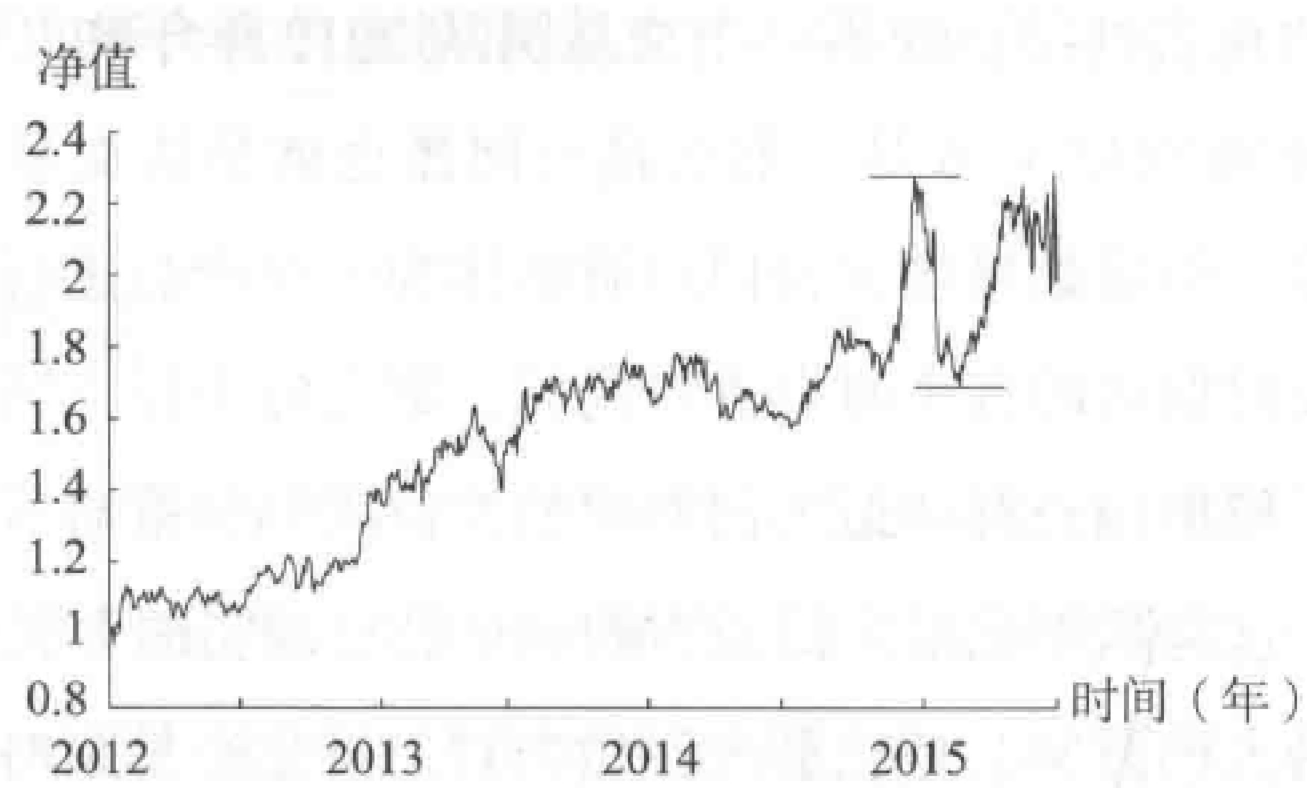


图 9-2 推进分析下自回归策略的净值走势与最大回撤

需要注意的是，最大回撤率这一指标存在一个特殊的属性，那就是其计算依赖于样本长度的选择。如果假设一个量化交易策略的风险程度不随时间的推移而发生变化，那么其 10 年的收益率标准差和 5 年的收益率标准差应该是相同的，但是 10 年的最大回撤率和 5 年的最大回撤率可能就不尽相同，前者大概率大于后者。为了使得风险度量不仅能够横向对比，同时也能够更直接地推广到对未来预期并保持一致，可以使用诸如年化最大回撤率之类的标准化度量形式。

年化最大回撤率的计算方式依赖于对最大回撤率的计算。仍然用第 7 章 7.4 节中的自回归策略举例，如图 9-3 所示，将原来完整的模拟交易期分隔为若干个以一年为长度的新样本期，即图中的一个灰色框，每个样本期较上一个向前推进一天。分别计算每个灰色框期间下的策略最大回撤率，最后将所有的最大回撤率综合并求取平均值，即为该策略的年化最大回撤率。经过这样的处理，使得最大回撤率这一指标也具有了由历史向未来推算时的普适性。

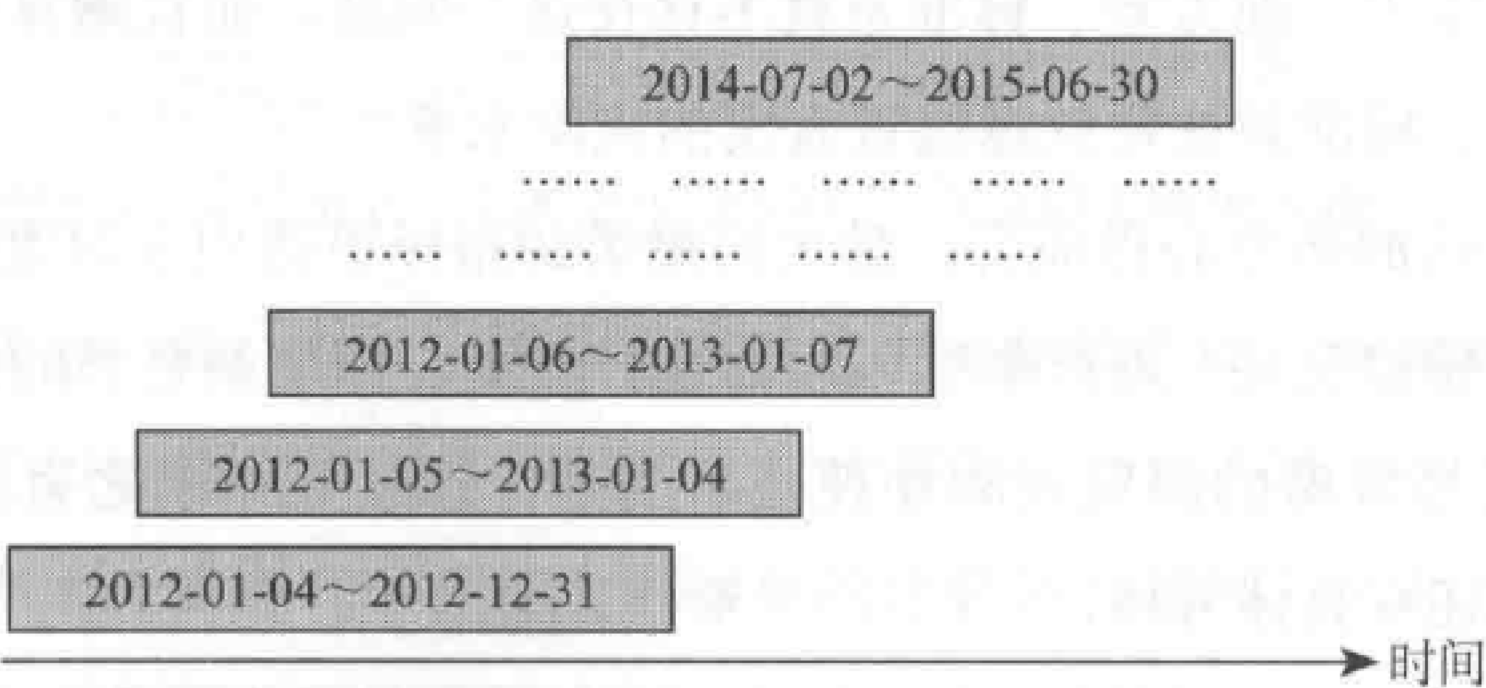


图 9-3 自回归策略年化最大回撤率的计算过程示意

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

从另一个角度看，最大回撤率与年化最大回撤率在设置上的区别也使得它们分别适用于不同的现实情况。如果一个交易员从 2012 年年初开始使用该自回归策略，并一直持续到 2015 年 6 月，那么最大回撤也就是该交易者在这段时期内所承受的最大风险。但是如果将该自回归策略作为一个产品进行售卖，而该产品的用户可能在这段时期内的各个时点进行购买，那么每个用户所承受的最大回撤也不尽相同，这时使用用户群体最大回撤率的均值来刻画策略所具有风险，就更有实际意义一些，这也是年化最大回撤率的一个贴合现实的含义。

可以看到，最大回撤率、年化最大回撤率对于量化交易策略风险的描述直接构筑在策略净值的基础上，接近实际表现，更为直观也更容易理解。因此许多实战派的交易员更喜欢使用这一指标，而且往往不仅用其来进行事后的风险描述，也会将其作为风险特征的度量加入策略模型中进行优化或者其他数学处理。

但是遗憾的是，最大回撤类指标的数学性质并不突出，会增加最优化等数学过程的复杂程度。与此同时，最大回撤类指标对数据样本选取的依赖程度较高，有可能因为变更了部分数据或者调整了样本期而产生较大的变化，因而更有可能产生过度拟合的问题。

相比较而言，方差、标准差就具有更好的数学性质，可以有效降低策略模型的求解难度。方差、标准差也更为稳定，由于其针对收益率时间序列做出统计，因此部分数据的变化对其最终结果的影响往往较小。基于这一类风险指标构造出的交易策略，在推进分析乃至实际使用中也因此会具有更好的稳定性。

还可以从另一个角度理解这种稳定性的差别。当我们根据样本数据测算出一段时间内的最大回撤水平之后，只要未来的策略运行时间足够长，最大回撤率一定会超过这一水平。而方差、标准差就不存在这个问题，而且测算时间越长，统计得出的方差、标准差往往就越接近真实的风险水平。

就作者自己的研发心得而言，最大回撤类的指标更适用于对策略交易（包括回溯测试下的模拟交易）进行事后风险描述，方差、标准差之类的指标则更适合在构建量化交易策略的研究过程中使用。当然，这也是一家之言，没有一定之规，具体问题还应具体分析。

在我们看来，研究人员使用某个风险度量的前提，应该是在逻辑和内涵上尽

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

可能地理解该风险度量，也只有熟悉了各种风险度量形式的含义和运用范围，才能根据具体的情况选择合适的度量指标，并且合理地使用它。

9.2 其他风险度量

最大亏损率是一个类似于最大回撤率的风险度量指标，同样偏向于对量化交易策略的实际表现进行描述。两者的差异在于，最大回撤率是一段时间内的任一时点向后推进，策略净值走到最低点时的净值回撤幅度的最大值，而最大亏损率则是从策略交易的起始时点向后推进，策略净值走到最低点时相对于起始点的净值回撤幅度。

再用推进分析下的自回归策略交易结果为例来说明最大亏损率并不直观，因为参考图 9-2 可以发现，实际上该策略的净值只在第四个交易日收盘时低于初始净值 1，而其他所有的时间内净值水平都是高于 1 的，最大亏损率接近于 0。为了使视觉效果更加明显，这里改用推进分析下的均线趋势策略为例进行说明。如图 9-4 所示，在三年半的模拟交易期内，策略的最大亏损率为 34.92%，最低点为 2014 年 4 月 9 日收盘时的 0.6508，图中使用一条横线进行标明。

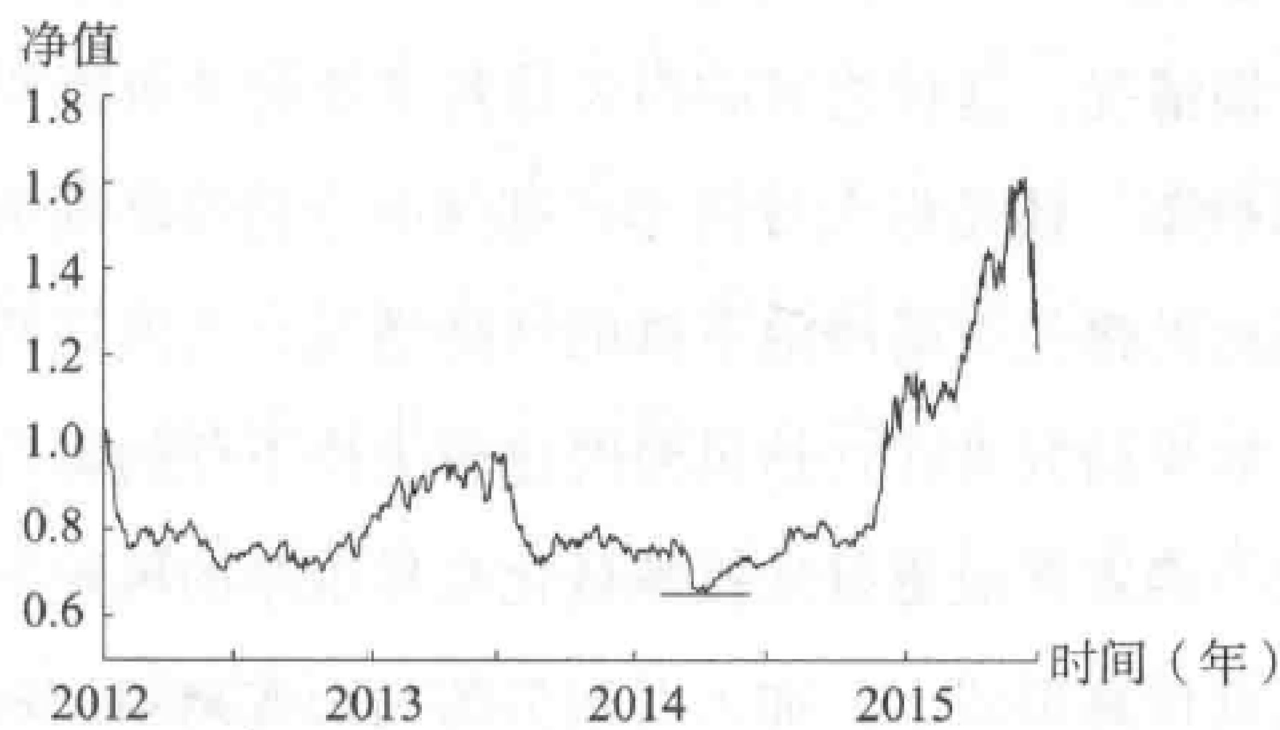


图 9-4 推进分析下均线趋势策略的净值走势与最大亏损率

同最大回撤率一样，最大亏损率也可以通过年化的方式，使其不仅能够横向对比，同时也可以在时间纵向上保持一致，方便研究者由历史到未来进行推广。年化最大亏损率的计算也与年化最大回撤率的计算框架相同，如图 9-5 所示，将原来完整的模拟交易期分隔为若干个以一年为长度的新样本期，即图中的一个灰

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

色框，每个样本期较上一个向前推进一天。分别计算每个灰色框期间下的策略最大亏损率，最后将所有的最大亏损率综合并求取平均值，即为该策略的年化最大亏损率。

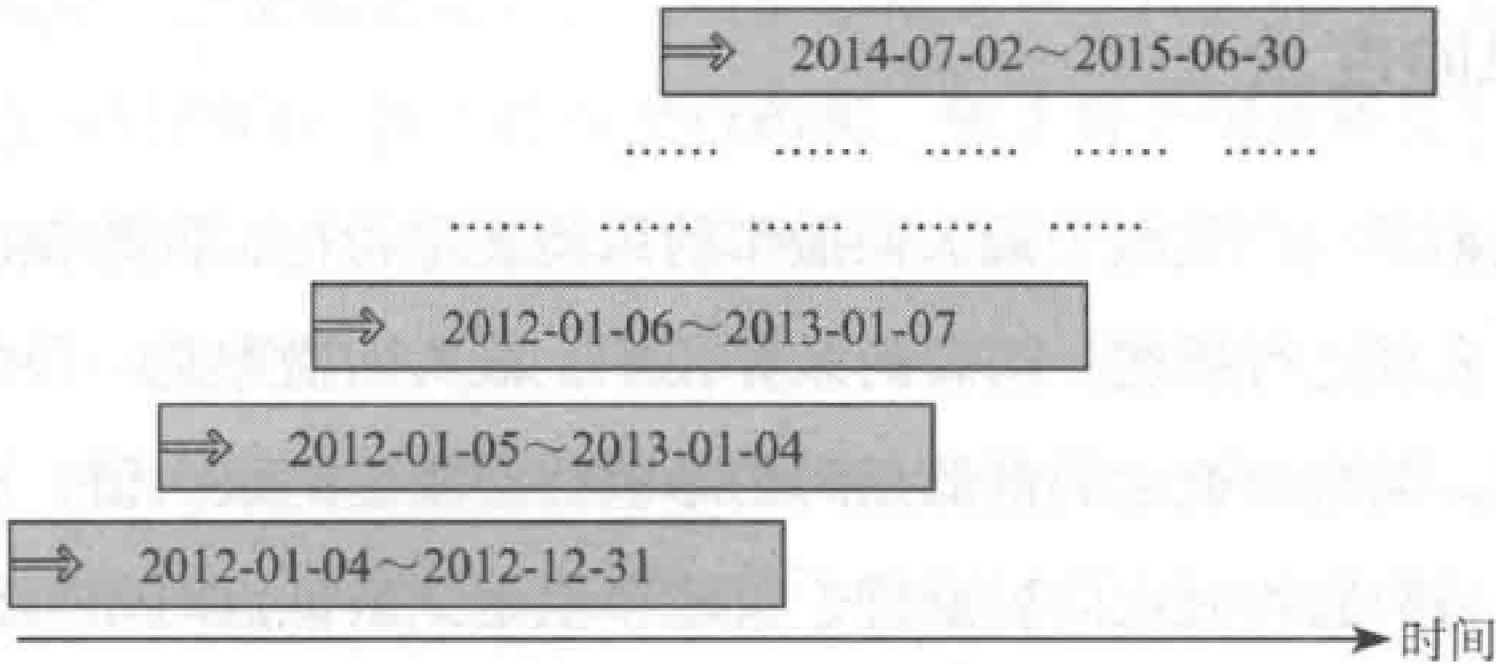


图 9-5　均线趋势策略年化最大亏损率的计算过程示意

由图 9-3 和图 9-5 可以看到，将最大回撤率与最大亏损率进行年化的计算方法都是相同的，其他一些指标的年化计算也可以参考这种做法进行处理，读者可以自行举一反三。而最大回撤率与最大亏损率的不同之处，主要在于锚定的参考点不同，最大回撤率是参考相对的净值高点进行计算，最大亏损率则是参考初始点进行计算。类比到实际操作中，最大回撤率的使用者更在意相对的损失，其把回撤之前所赚的资金视为自己的资金进行对待，而最大亏损率的使用者则更在意初始本金之上的亏损情况，这种绝对的损失是其主要的考量因素。

不论是最大回撤率、还是最大亏损率，都倾向于将策略损失的可能性作为风险的主要来源，刻画的都是策略净值下跌的特征情况。当然这两者都具有数学易用性较差的缺点。如果研究者在评估风险时比较关注下跌的影响，认为损失的不确定性比之盈利的不确定性能更好地反映量化交易策略的风险，那么其实可以在方差、标准差的基础计算形式上，加入对损失的判别或刻画，从而既在一定程度上保留风险指标的良好数学特性，同时也更好地反映策略的实际表现。

这种关注于损失的方差形式，在数学上一般被称为半方差，定义与普通的方差类似，唯一的区别在于，计算半方差时，整个统计过程中只使用那些小于均值的样本。设推进分析中模拟交易期的收益率均值为 μ，收益率小于均值 μ 的交易日共 N_S 天，其中第 i 个小于均值 μ 的策略收益率为 r_i^S，则模拟交易样本下的半

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

方差$\bar{\sigma}^2$可以计算如下

$$\bar{\sigma}^2 = \frac{1}{N_S - 1}\sum_{i=1}^{N_S}\left(r_i^S - \mu\right)^2$$

同样的，为了与收益率的量纲保持一致，可以在样本半方差的基础上进行开平方处理，得到半侧的标准差计算方法如下

$$\bar{\sigma} = \sqrt{\frac{1}{N_S - 1}\sum_{i=1}^{N_S}\left(r_i^S - \mu\right)^2}$$

可以看到，半方差实际上度量的是较差的收益率距离稳定收益的偏离度。而当收益率均值μ接近于零时，半方差实际上度量的就是策略损失情况的不确定性。同时，半方差还存在另外一种形式，也就是所谓的目标半方差，其将样本选取的阈值由均值μ变更为某一具体的收益率目标。如果我们设定该收益率目标为零，那目标半方差度量的就是策略损失的不确定性了。

但是值得注意的是，半方差这一类的风险指标由于引入了对损失的刻画，因此复杂程度还是要高于方差、标准差的，数学性质上也要逊色不少，在一些量化交易策略的模型中应用起来也更为困难。例如在马科维茨的最优投资组合理论中，基于多个资产的方差计算投资组合方差较为方便，但是基于多个资产的半方差计算投资组合的半方差就需要复杂得多的数学处理了。

最大回撤率、最大损失率、半方差、目标半方差这些风险度量指标都侧重于刻画量化交易策略损失一侧的风险情况，我们可以将它们归为一个大的风险指标类别，一般称之为下侧风险或者下行风险。在这一类风险指标中，还包括另外一些关注损失或其他负面情况的度量方式，例如在金融实务中有着广泛应用的风险价值（Value at Risk，VaR）。

风险价值代表的是策略在某个概率下可能亏损的程度，或者更宽泛一点说，策略极端亏损的程度。其数学上的定义是，给定时间长度范围，设定某一个概率P%，风险价值R可以使得策略收益小于$-R$的概率恰好为P%。实际上，$-R$也就是收益的概率分布下P%的分位数。如图9-6所示，给定P%，就可以通过策略收益率的概率分布计算得出风险价值R。举例而言，给定5%的概率，如果5%

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

概率下的风险价值为 0.8，则意味着量化交易策略的损失大于 0.8 的概率为 5%。

预期损失是由风险价值引申出的风险度量指标，也被称为条件风险价值。在定义上，其是指某一个概率 P% 下，所有可能的损失情况的期望值，也就是图 9-6 中灰色区域的期望值。预期损失在数学性质上具有风险价值所没有的次可加性，因此在研究投资组合风险时，有时候会是一个比风险价值更为合适的选择。

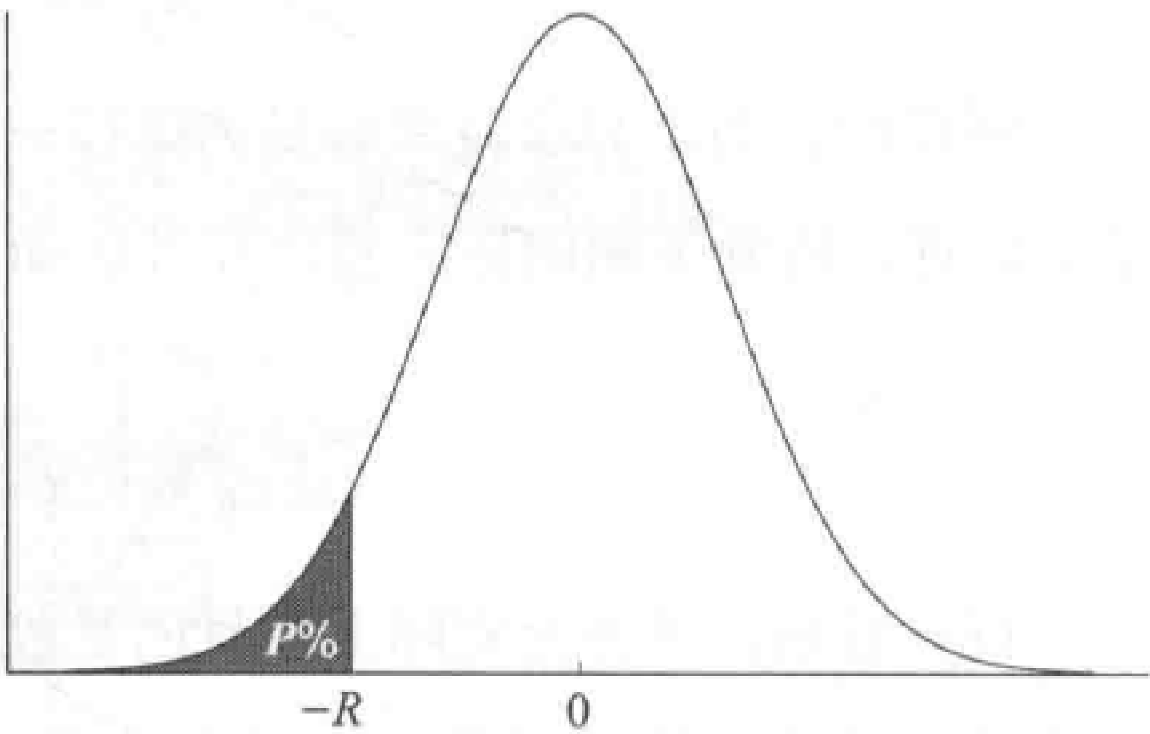


图 9-6　风险价值和损失概率的计算示意图

损失概率是一个与风险价值在内涵上同源的风险度量指标。区别在于，风险价值先设定概率 P%，后计算分位数 −R，而损失概率则是在设定好分位数 −R 的基础上，计算该分位数所对应的概率 P%。也就是说，损失概率代表的是策略收益小于某一个设定值的概率情况，其衡量的是量化交易策略极端亏损的可能性。

风险价值、预期损失和损失概率讨论的都是极端情况下的策略损失风险，但是其刻画的风险又不如最大回撤率、最大损失率所刻画得那么极端，更偏向于对尾部某一个区间内的极端情况进行描述。因此，可以将它们看作介于最大回撤率和半方差之间的一类下侧风险指标。

纵观前面提到的收益率概率分布、方差、标准差、最大回撤率、年化最大回撤率、最大亏损率、年化最大亏损率、半方差、目标半方差、风险价值、预期损失、损失概率等，都是基于历史数据得到的统计指标，它们更倾向于对量化交易策略在已知数据下的表现进行描述。因此，将其用来对量化交易策略实际运行的风险情况进行事后度量和表述，是恰当合理的。

但是在研发量化交易策略时，它们更重要的作用往往是作为策略风险的表征加入策略模型构建当中。在这种使用方式下，实际上暗含着一个假设，即样本时期策略表现出来的风险与策略在未来的风险是一致的。建立这样的假设之后，这些对于风险表现的描述才可以看作是量化交易策略所固有的风险，从而将其应用于策略的构建研发等方面。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

在量化交易策略度量风险时，存在另外一种假设，即风险变化的模式是不变的，或者说在样本时期和未来，风险模型是保持一致的。基于这种假设，可以引入其他一些风险度量的方法，如条件方差等。一个计算条件方差比较著名的模型是广义自回归条件异方差（GARCH）模型，其假设某个时间点 t 的收益率 r_t 可以分解为预期收益率 $E(r_t)$ 和残差项 ε_t

$$r_t = E(r_t) + \varepsilon_t$$

然后假设收益率条件方差 h_t 和往期残差项 ε_{t-i} 以及往期条件方差 h_{t-i} 相关

$$h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i h_{t-i}$$

式中，p 和 q 是存在影响的往期期数，常见的 GARCH 模型一般将这两个值都取为 1。而当预期收益率 $E(r_t)$ 为零时，可以将样本所显示的收益率直接作为残差项代入条件方差模型，简化设置。

在确定方差变化的模式之后，通过在样本数据上进行拟合估计就可以得到相应的参数值，进而得到具体的风险运行情况。如果读者们对具体的估计过程感兴趣，可以自行参阅相关的计量经济学资料。

这两种关于风险的假设比较起来，一个是假设整体风险保持不变，另一个是假设风险变化的模式保持不变，前者偏静态，而后者偏动态。在第 1 章 1.2 节阐述量化交易策略优缺点时曾经提到过，由于数量化方法需要一定量的数据样本进行研究，而相应的数据都是随着时间的推移逐渐产生的，因此从数据中抽取的数量化特征也只会随着时间的推移逐渐变化，这种缓慢变化的特性会导致量化交易策略较难适应市场的转折。动态的风险度量方法在这个问题上稍有优势，因此一些追求时效性的量化交易模型会倾向于使用这一类的方法来对风险状态进行刻画和调整⊖。

在这两节中，我们介绍了若干种风险度量的指标和方法。可以看到，各种风险度量方式都具有自身的特性，刻画的风险特征不同、数学特性不同、模型易用性不同、所基于的假设也不尽相同。这里再次强调，在量化交易策略的研发过程

⊖ 例如 BARRA 的风险模型 USE3 等，就使用 GARCH 模型所得到的市场风险来动态调整静态模型下的风险程度。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

中，研究人员有必要熟悉各种风险度量形式的含义和运用范围、区分所使用的风险度量的适应环境，从而保证在不同情况下使用合适的风险度量方法，合理化整个研究流程。

9.3 风险和收益的结合

前面提到过，策略或资产的收益本身是有益的、正面的，但是收益的不确定性或者说风险则相反，当风险程度增加时，量化交易策略的最终结果会呈现出不稳定的表现，这种不稳定性也就是风险所带来的负面效应。收益和风险作为量化交易策略研发中两个最为重要的特征，一个正面一个负面，既可以作为我们事后评判策略表现的参考指标，也可以作为事前的度量工具参与策略模型的构建工作。

如第 2 章中所述，在基本的策略研发流程下，策略或资产的收益情况，是与量化交易策略的买卖行为直接相关的。即基于买点、卖点的选择，使得策略得到一个正的整体收益，或是通过判断交易资产的未来收益正负，来决定买入或卖出资产。而风险则会影响到量化交易策略中交易仓位的设置，也在一定程度上与收益因素共同决定了策略的买卖设置。

而在图 2-4 所示的量化交易策略研发流程中，我们将策略模型所使用的目标函数进行了唯一化，即策略的风险调整后收益，也就是风险与收益的一种有效结合。通过考察这一结合指标，买卖和仓位的设置可以同时作为参数进行优化，进而得到最终的量化交易策略。

在这里，风险调整后收益是一种由收益和风险共同组成的数学指标，其随着收益的增大而增大，随着风险的增大而减小。而由于收益是正面因素，风险是负面因素，因此风险调整后收益就可以直观表现出量化交易策略的优劣。将风险调整后收益作为唯一的目标函数也就能够让策略优化过程更为直接。

当然，将风险调整后收益作为策略表现的一个事后度量指标也是有所裨益的。例如，存在两个待选的量化交易策略——策略 A 和策略 B，策略 A 的收益更高但风险也更高，而策略 B 的收益较低但风险也较低时，使用分离的收益度量指标和风险度量指标就很难判断出哪个策略更好。而通过比较两个策略的风险

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

调整后收益，就可以准确地判断出哪个策略更加优秀。

准确地讲，风险调整后收益是一类指标的统称，只要是正面度量收益、负面度量风险的数学形式都可以算作风险调整后收益。夏普比率是最为著名也是应用最为广泛的风险调整后收益指标，以其创造者夏普的名字命名。夏普比率度量的是交易策略或者投资资产每承受一单位的总风险，可以产生多少的超额收益。具体计算形式为

$$夏普比率=\frac{R-r_f}{\sigma}$$

式中，R 是交易策略或者投资资产的收益率；r_f 是无风险收益率；σ 则是交易策略或者投资资产的收益率标准差。

信息比率是另一个经常被用到的风险调整后收益指标，特别是在研究量化选股策略时，研究人员往往更偏向于使用信息比率。实际上，信息比率和夏普比率的构造方法十分相似，度量的同样是交易策略或者投资资产每承受一单位的总风险，会产生多少的超额收益。两者最大的差别在于分子中被减去的，一个是基准收益率，一个是无风险收益率，具体如下

$$信息比率=\frac{R-r_B}{\sigma(R-r_B)}$$

式中，R 是交易策略或者投资资产的收益率；r_B 是基准收益率；σ（$R-r_B$）则是交易策略或者投资资产的收益率减去基准收益率后，所得到的新收益率序列的标准差。

这里所说的基准收益率，是由研究人员选取一个合适的基准资产、基准策略或基准指数所具有的收益率，其与所研究的交易策略或投资资产具有某些相同的特性，因此可以用来进行对比。观察夏普比率的计算形式可以发现，其实夏普比率就是将比较基准选为无风险资产的信息比率。

前面若干个章节讨论选股策略的实际表现时，我们都将同期的沪深300指数走势加入图中，用来与选股策略进行对比，同时也考察了逐月地将选股策略得到的股票组合收益率减去沪深300指数的收益率后，所形成的新的收益率序列及其净值走势。原因在于量化选股策略总是做多股票，因此投资组合中总是包含有市

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

场组合成分以及相应的市场系统风险。这正是研究量化选股策略时经常使用信息比率的原因，通过减去市场组合基准的收益率，能够更好地反映出策略在选股方面的真实能力。

什么时候更适合使用夏普比率，什么时候更适合使用信息比率，需要研究人员理解量化交易策略或是其研究对象的收益组成、风险组成以及各个组成部分的具体性质。例如第 4 章和第 7 章所涉及的择时策略，多空灵活，更适合将无风险资产进行对比，那么夏普比率就是一个较为合适的风险调整后收益指标。不过，在实际使用中，许多研究者往往会混用夏普比率和信息比率，读者如果遇到这种情况还需注意。

在风险调整后收益指标中，与夏普比率和信息比率类似的，还包括索提诺比率。其分子端的构成和夏普比率相同，分母端的标准差则换为半侧标准差$\bar{\sigma}$。也就是说，该指标中所包含的风险项刻画的是交易策略或者投资资产的下侧风险，如下：

$$\text{索提诺比率} = \frac{R - r_f}{\bar{\sigma}}$$

式中，$R-r_f$也可以在某些情形下变化为$R-r_B$，从而加入适合的基准进行对比。

以上三种风险调整后收益指标都是通过收益除以风险得到，因此也具有一些较为相似的性质。例如，当收益和风险都是由日度数据计算得到的时候，是不能与由月度数据计算得出的结果直接进行对比的。不同时间频率的样本数据，会使得这种除法导向的指标的度量基准不同。

如果要横向对比两种不同样本频率下的量化交易策略，例如前文中使用日度数据进行研究的择时策略和使用月度数据进行研究的选股策略，那么需要将这些风险调整后收益指标调整到一个频率基础之下，如将指标年化。而进行年化处理时，往往是分别将收益和风险年化，然后做除法得到。收益的年化相对而言比较好理解，标准差这类指标的年化则一般是乘上频率比例的方差，例如月度数据下的标准差年化，即在其数据上乘以 $\sqrt{12}$ 得到。

除了上述三种指标之外，还存在其他一些由除法得到的风险调整后收益指标。实际上，除法导向是一个非常好的收益风险组合框架，分子端的收益刻画正面效果，分母端的风险刻画负面效果，符合研判直觉。广义上而言，只要找到合

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

适的收益度量方式和合适的风险度量方式，将其通过除法结合在一起就是一个可用的风险调整后收益指标。另外，做除法时保持两者的量纲一致是较为推荐的，可以使指标具备较强的逻辑内涵和较好的数学特性。

例如作者较为偏好的一个做法，用年化收益率除以年化最大回撤来表示量化交易策略的风险调整后收益。由于最大回撤是一个更偏向于对策略实际表现进行描述的风险度量形式，因此这一风险调整后收益也更适合用来对量化交易策略的表现进行事后评判。将其作为目标函数进行优化也是可行的，不过该指标不太良好的数学特性会引发一些附带的问题，因此也建议其他研究者谨慎为之。

另外例如特雷诺比率，也是一个通过除法计算得到的风险调整后收益指标。其计算形式如下

$$\text{特雷诺比率} = \frac{R - r_f}{\beta}$$

可以发现，其与夏普比率的区别，同样在于风险度量方式选取的不同。这里的 β 代表策略或资产所具有的系统风险，这一概念来自第 1 章曾提到过的资本资产定价模型，也就是策略净值随市场资产同向运动所形成的没有风险补偿的部分风险。

在对收益的风险调整过程中，收益与风险的搭配方法不止除法一种。前面已经提到过，只要是对收益给予奖励、对风险给予惩罚的算法，都可以作为风险调整后收益的度量指标使用。一个更简单的做法是用收益减去风险，例如用收益率减去方差

$$\text{风险调整后收益指标} = R - \theta \cdot \sigma^2$$

式中，θ 是人为设定的比例系数，用来调节对风险的惩罚力度，在一定程度上表现了指标使用者的风险偏好。

第 1 章曾提到过的马科维茨最优投资组合理论，其基本计算形式是在方差一定的基础上最大化收益率，或是在收益率一定的基础上最小化方差。而作为一种变种形式，我们也可以将收益率减去方差作为唯一的优化目标，通过最大化这一指标，在设置好惩罚系数 θ 的基础上得到最优的投资组合配置。

一般而言，将风险和收益相结合的风险调整后收益指标，不太会使用类似于 GARCH 这样的动态调整模型。就算存在，也只是将动态调整的收益和动态调整

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

的风险通过除法或减法结合起来，用作事后的量化交易策略绩效度量，其优点在于更具有时效性，更能适应于市场或策略的转变。但是若将这种动态的风险调整后收益指标用于策略研发过程，往往会使得策略模型过于复杂，得不偿失。

9.4 止损

将止损的内容放置在哪一个章节中，是让作者在构思整本书时深受困扰的一个问题。止损本身是一个交易行为，在量化交易策略的组成成分中与买卖行为的性质等同。但是与买卖行为不同的是，止损的目的主要是为了控制风险，而买卖的目的主要是为了盈利，或者说获得更好的风险调整后收益。

同时，止损是一个偏向于人为主观确定的交易行为。在全书开头对量化交易策略给出定义时，作者曾说明，在构建交易策略的过程中，允许存在定性的或者人为主观的成分。止损就是一个在策略构建过程中存在较多定性成分的组成部分，虽然研究人员完全可以将其作为一种买卖决策加入整体模型从而进行优化处理，但是实际的止损设置往往还是与量化交易策略使用者的风险规避程度具有更为密切的关系。

当然，止损在量化交易策略构建完成之后，是具有明确的数量化规则的，可以用来无偏差地指导交易决策，因此止损本身完全符合量化交易策略的定义。综上所述，对止损而言，在量化交易策略研发框架中进行介绍、将其加入买卖决策部分的论述中或是在风险体系内加以介绍，都是可行的。

然而，让作者感到抱歉的是，在本书这种介绍量化交易策略研发流程的目的背景下，止损的加入似乎会影响到整体论述的明确性。一方面，止损与买卖行为的交易性质类似，在策略研发框架和买卖决策部分加入对止损的介绍会使得理解较为困难；另一方面，止损本身较强的定性特质会让全书的量化基调受到影响；此外，在实际策略的例子特别是量化选股的例子当中加入止损的设置会极大地复杂化处理过程，希望保持本书简单易读的特性也是作者不愿意增加篇幅详述止损的一个原因。

但是我们需要认识到，止损在量化交易策略研究与量化交易实际应用中具有

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

十分重要的地位，一些交易员和研究者也把止损视作风险管理的重要工具[⊖]。在一本介绍量化交易策略的书籍当中绝口不提止损是一个有失偏颇的处理方法，因此作者选择在本章中对止损进行简单的介绍。如果读者对这一风险控制工具感兴趣，可以参考其他的交易资料，并尝试着学习如何在实践中将止损设置融入整体的量化交易策略研发。

一般来说，量化交易策略中的止损包括两类。第一类针对单次交易，如果净值因为这一笔交易而损失到了一定程度，则不必等到相应的退出点出现，直接清空仓位止住损失。在这样的交易中，止损代替了退出买卖行为，终止了一次完整的交易，与买卖行为的性质是一样的。

图9-7说明了这一类止损在量化交易策略研发流程中的位置。在图9-7中，追求收益的目标驱动了买卖决策的形成。同时，对单次交易风险的控制，也通过止损的形式，影响了买卖决策的结果，这一过程就是图中的虚线箭头。

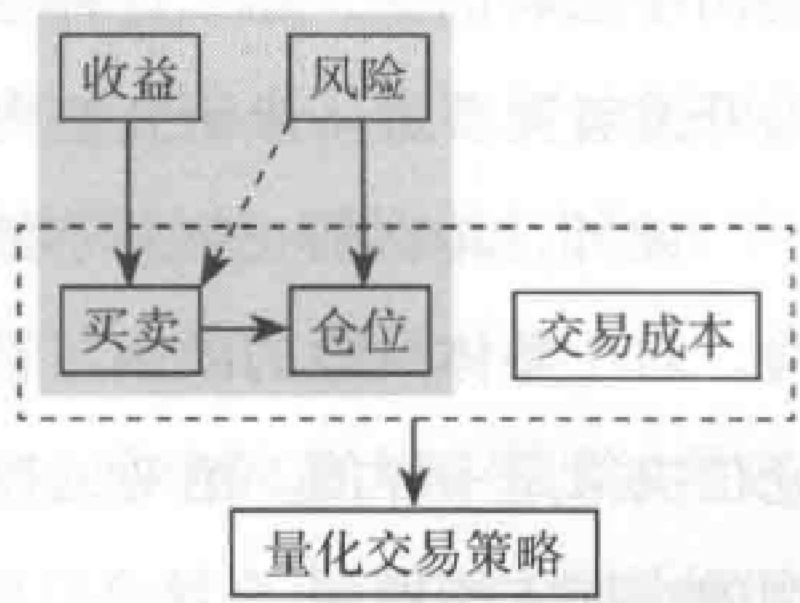


图9-7 止损在量化交易策略研发流程中的位置

第二类是针对整个量化交易策略，如果实际运行一个量化交易策略，使得净值从开始运行时到现在损失到了一定程度，则终止该策略的运行，判定该量化交易策略失效，并放弃使用。这在一定程度上可以看作最大损失率的一个实际应用，即设定一个值，当最大损失率超过该值时，就退出量化交易策略的运行。

第二类止损更偏向于量化交易策略的执行环节，存在着更多的人为确定成分。其具体的设置有时候与策略使用者的风险偏好有关，有时候与相应投资产品的设计有关。例如一些基金产品的合同中会设计风险警戒线，这种合同上规定的亏损额度往往就是这一类止损设定的重要来源。也正是因为人为因素过大，大多数量化交易策略的研究资料并不讨论这一类的止损行为。

⊖ 杰西·利弗莫尔（Jesse Livermore）有过“那些买进或卖出后就出现浮亏的东西说明你正在犯错，一般情况下，如果三天之内依然没有改善，立马抛掉它”的论述。海龟交易法则中也有专门的止损设计。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

第一类止损是我们在设计量化交易策略时主要研究的止损方式，即在一次交易的建仓后，如果资产走势并没有按照我们希望看到的方向运行，进而产生了净值回撤，那么当回撤达到多大程度时，就应该认定这一次的建仓是错误的，清空仓位止损。

这样的止损只是主动退出某一次交易的处理方法，并不影响交易者按照量化交易策略的信号进行下一次的开仓与买卖。因此完全可以将其看作整个量化交易策略的一部分，并设置具体的参数加入建模与优化的过程。

但是也正因为这类止损止住的只是一次交易的损失，所以止损对于风险的控制只体现在单次交易之上。通过单次交易的风险来评估量化交易策略的整体风险仍然是一个待处理的工作。例如，如果出现连续止损的情况，净值的回撤很可能会出乎意料的大，那么这种连续判断错误、连续止损的出现概率有多大，就是策略开发者需要进一步研究的内容。

此外，单次的止损并不能一劳永逸地帮助我们限制好量化交易策略本身的风险。对于整体风险的限制与管理，其只有相对的作用，而没有绝对的作用，这与仓位决策是一样的。绝对的控制风险，只有前面提到的第二类止损，即完全退出型的止损才能做到。

以上是对止损行为的一些简单讨论，希望读者能通过这些论述对止损形成一个基本的认识。本节最后介绍第一类止损当中几种较为常用的形式，希望聊胜于无，算是完成了全书对止损的涵盖。

第一种是固定比例止损，这也是最为常见的一种止损方法。具体的形式为，以一次交易开始时的净值为基础，当净值下跌到初始净值的一定比例时，执行止损，清空原有仓位。以最简单的满仓持有股票为例，设定止损比例为 20%。假定在一只股票 100 元的时候持有，出现卖出信号之前，如果股票价格下跌到 80 元，则卖出所有股票，即为止损。

第二种是固定金额止损。具体形式是，以一次交易开始时的净值为基础，当净值下跌到一定的数额时，执行止损，清空原有仓位。仍然以满仓持有股票为例，设定损失金额为 200 万元。假定在一只股票 100 元的时候买入 10 万股，此时的资金总量为 1 000 万元，那么在出现卖出信号之前，如果股票价格下跌到 80 元，资

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

金量变为 800 万元，则卖出所有股票，即为止损。而如果买入股票时的资金量为 2000 万元，即买入 20 万股，则在股票价格下跌到 90 元时就要执行止损了。

对比这两种比较常见的止损方式，作者更倾向于固定比例止损，同时也建议读者更多地关注这种止损形式。原因在于，虽然固定金额止损在实际操作中更为直观，但是其并不能很好地适应资金净值的变化。随着量化交易策略的持续运行，初始设计的固定金额止损可能会越来越不适用于新的净值情况，导致风险暴露变得过大或过小。而固定比例止损就不存在这一问题，通过使用比例这一数学工具，止损过滤掉的风险大小可以始终与当前的资金量保持一致。

还有一种止损形式是通道突破止损。举例而言，从某一次交易的建仓为始，如果资产价格低于 20 日内的最低价，则清空仓位止住损失。从描述中我们也可以发现，其实这种止损形式与一般的买卖点判断形式并无区别，完全可以看作在原有退出策略的基础上又增加了一种退出判断。只不过在形成退出机制的同时，这样的操作同时具有止住损失的特性罢了。还有很多其他技术性的退出判断方式，其实也都具有同样的特性。

综合来看，这一类止损的目的，都在于规避单次交易中的重大损失。如果遇到一波巨大的单边趋势性走势，往往一次看似普通的开仓就会由于方向判断相反而损失掉大量的资金。通过加入单次交易下的止损，量化交易策略的使用者就可以防范这种突然而来的黑天鹅，避免被一次大的波动所击倒。

不过，当黑天鹅不是以一次大波动的形式出现，而是连续地与交易建仓唱反调时，针对单次交易的止损操作就无能为力了。更为极端的情况是，交易员刚刚止损，行情又开始顺着建仓的方向运行，策略连续被止损、连续亏损，这种情形是十分打击量化交易策略使用者对止损设置本身的信心的。

当然，作者并不是想说止损是一个不好的风险管理方式。实际上，在作者的理解中，这是一个仅次于仓位决策的风险控制行为，虽然在广度等方面不及仓位决策，但是更为直接、在操作上更接近下侧风险控制的实质。

研究者们应该理性地认识到，量化交易策略中的每一个组成部分、每一种操作都是具有其自身特性的，如何发挥优势、回避劣势、合理地融合到策略整体当中去，才是一个合格的研究者应该重点考量的问题。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

第10章

仓位决策

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

10.1 凯利公式

凯利公式是最为著名的仓位决策方法，由约翰·凯利（John Kelly）于 1956 年创造发表[1]，其后在博彩和交易领域中被广泛地应用。其是一个在特定的赌局中，用来计算出每次游戏中应投注的资金比例，从而使得重复下注后的资金增长期望值达到最大的算法。而投资交易往往与重复性赌博具有很大程度的相似性，因此扩展了公式的适用范围。

与其他相关资料类似，本书也从一个易于理解的例子开始对凯利公式进行介绍。假定存在一个抛硬币的简单赌局，赌客押 1 元，如果抛到正面押金退还，并赔给赌客 2 元；如果抛到反面，则赌客输掉押金 1 元。在硬币和其他庄家操作环节没有问题的情况下，硬币抛到正面或反面的概率一样，也就是赌客的胜率为 50%，但是赔率为 2，这是一个对赌客有利的赌局[2]。

如果赌客可以一直持续地进行这样的赌博，并且可以在每一次赌局中押注任意的数额[3]，那么赌客需要决策的就是押注多少才最为有利。假设某位赌客初始的资金量为 1 元，那每一次都全额押上肯定不是一个正确的选择，因为就算运气再好，也总会遇上一次反面，本钱输光就出局了。对于一个对赌客有利的赌局而言，这样的押注方法无疑等于没有头脑。

固定数额的押注也不那么正确。例如每次押 0.1 元，如果运气不好，可能若干次赌局之后，赌客会随着押入最后 0.1 元并抛出反面而输光赌金出局。实际上，由于赌局在概率上对赌客有利，因此赌客最好保证自己能够一直赌下去，不能让自己陷入输光本金的可能性当中。

在这种前提下，固定比例的押注就成为一个比较好的解决办法。可以类比第 9 章 9.4 节介绍止损时所述，固定金额虽然在实际操作中更为直观，但是其并不能很好地适应资金净值的变化，而固定比例的设计可以始终与当前的资金量保持一致。最为重要的是，每一次硬币抛出反面的时候，赌客损失掉的资金都只占

[1] Kelly, J L. A New Interpretation of Information Rate. Bell System Technical Journal. 1956, 35 (4): 917–926.

[2] 在一对一赌博中，在概率上对赌客有利的赌局是很难出现的，这里仅作举例用。

[3] 包括小数金额，例如 0.000001 元。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

全部资金的一定比例，从而确保了赌客始终有数额为正的资金将赌局继续进行下去，永不出局。

那么现在，赌客需要决策的问题就变成了每次押注多大比例的资金最为有利。按照一般的赌徒直觉，押得大、风险大，收益就大，那似乎只要不是一次全押，比例越高应该盈利越多。遗憾的是，概率世界中这样的情景是不成立的。我们可以先简单看一下将押注比例设为 80%、50% 和 20% 时，赌客资金量随赌博次数增加而变化的具体情况[⊖]，如图 10-1 所示。

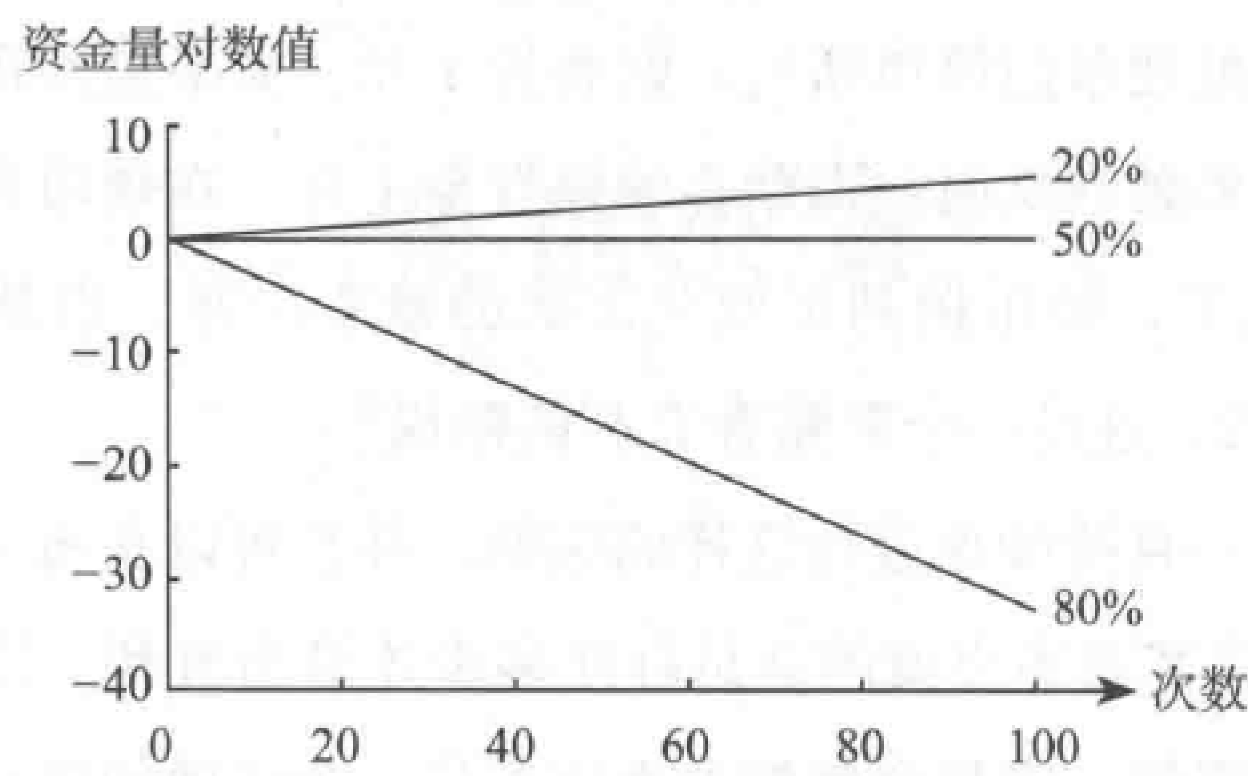


图 10-1 不同押注比例下资金量随赌博次数增加而变化的情况

图 10-1 中横轴为赌博的次数，纵轴为各次赌局下资金量期望的对数值。由图 10-1 中我们可以看到，当每次的押注比例设置为 80% 时，虽然赌客不至于一把输光本金，但是 100 次赌局之后其资金量也极度接近于零了。当押注比例设置为 50% 时，资金量的期望一直维持不变，说明这是一个赌客与庄家都不胜不败的局面。反倒是押注比例设置为 20% 时，资金量一直随着赌局的进行而持续增加。

这样看来，最佳的押注比例似乎是在 50% 以下的某一个值。我们不妨将 100 次赌局之后资金量期望随押注比例变化而变化的趋势情况展现出来。由于该例子中不涉及借贷与杠杆，因此押注比例的变化范围限定为 0 ～ 100%。这里纵轴中的资金量期望保持原数值，不再做对数变换，具体情况如图 10-2 所示。

⊖ 为保证曲线平滑，资金量的数值是各种情况下资金量的期望均值。同时取资金量的对数值，使得三条线在趋势不变的前提下变化幅度减小，方便展现于同一张图中。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

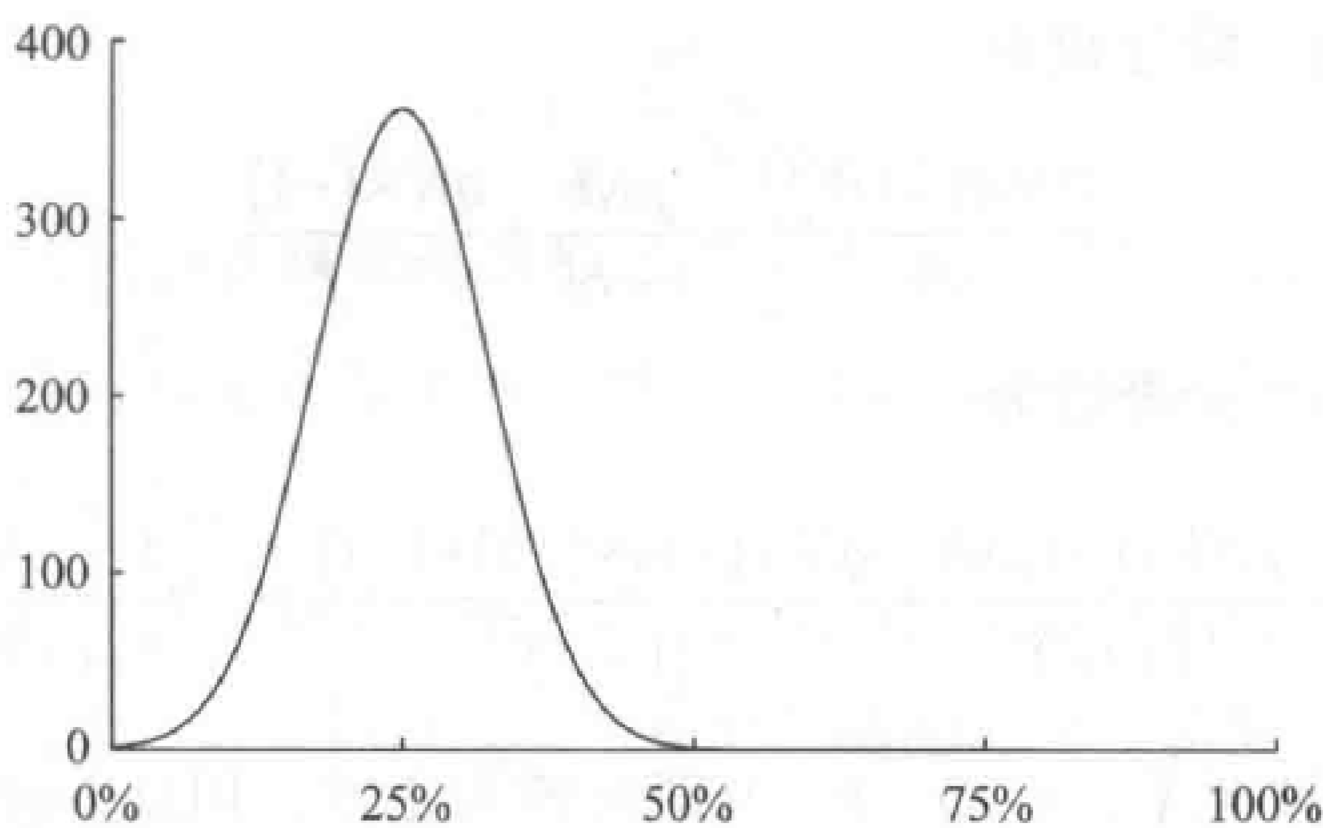


图 10-2　押注比例与资金量的趋势变化关系

从图 10-2 中我们可以发现，盈利水平由 0 的押注比例设置开始逐渐提升，到 25% 达到最大值，是初始水平的 361 倍。随后盈利水平开始随着押注比例的增加而不断下降，到 50% 维持不胜不败，比例大于 50% 的押注设置则会让赌客逐渐损失本金。那么很明显，在上面的赌局例子中，赌客的最优决策就是每一局都押注当时资金量的 25%，并持续性地玩下去。

接下来，我们将这个赌博的例子一般化，构造出一个数学模型来解释并推导凯利公式的基本形式。假设存在一个赌局，赌客可以一直持续地进行赌博，并且可以在每一次赌局中押注任意的数额。对赌客而言，其押中的概率为 p，即胜率为 p，押错的概率为 q。[⊖]当赌客押中时，押金退还，并赔给赌客押金的 b 倍；如果赌客押错，则输掉押金，也就是说，该赌局的赔率为 b。

如果赌客采用按固定比例下注的押注手法，每一次的押注金额都是当时金额的 f，f 为某一比例值。当赌局进行了 N 次之后，赌客资金量的期望值应为

$$E(W)=(1+bf)^{pN}\cdot(1-f)^{qN}$$

从数学意义上来说，赌客的决策就在于选择合适的 f，使得 $E(W)$ 最大。在优化之前，我们先讨论 $E(W)$ 的数学性质。先将 $E(W)$ 对数化，从而在保留 $E(W)$ 原有趋势特性的基础上更容易进行数学处理，最大化 $E(W)$ 实际上与最大化 $\log(E(W))$ 等价。$\log(E(W))$ 的形式可以描述为

$$\log(E(W))=pN\log(1+bf)+qN\log(1-f)$$

⊖ p 和 q 都应该为 0 ～ 100%，且满足 $p+q=100\%$

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

log(*E*(*W*)) 的一阶导数为

$$\frac{\partial \log\left(E\left(W\right)\right)}{\partial f}=\frac{pNb}{1+bf}+\frac{qN\times\left(-1\right)}{1-f}$$

log(*E*(*W*)) 的二阶导数为

$$\frac{\partial^2 \log\left(E\left(W\right)\right)}{\partial f^2}=\frac{pNb\times\left(-1\right)\times b}{\left(1+bf\right)^2}+\frac{qN\times\left(-1\right)\times\left(-1\right)\times\left(-1\right)}{\left(1-f\right)^2}=-N\left[\frac{pb^2}{\left(1+bf\right)^2}+\frac{q}{\left(1-f\right)^2}\right]$$

由于在实际情况下，p、q、b、N 都应该为正数，因此 log(*E*(*W*)) 的二阶导数是严格小于零的，这种特性使得 log(*E*(*W*)) 随 f 变化的形状与图 10-2 中的曲线形状类似，都是先上升后下降，存在一个总体上的最高点。

而这个最高点的 f 取值，应该满足 log(*E*(*W*)) 的一阶导数为零，也就是这一点上的切线应该平行于横坐标，可以参考图 10-2 的图形进行理解。通过求解 log(*E*(*W*)) 的一阶导数等于零这一方程

$$\frac{pNb}{1+bf}+\frac{qN\times\left(-1\right)}{1-f}=0$$

可以得到

$$f=\frac{bp-q}{b}$$

具体推导过程如下：

现有
$$\frac{pNb}{1+bf}+\frac{qN\times(-1)}{1-f}=0$$

则
$$\frac{pb}{1+bf}=\frac{q}{1-f}$$

$$pb(1-f)=q(1+bf)$$

$$(qb+pb)\,f=pb-q$$

由于
$$p+q=100\%$$ [⊖]

则
$$bf=pb-q$$

⊖ p 和 q 都应该为 0 ~ 100%，且满足 $p+q=100\%$

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

$$f=\frac{pb-q}{b}$$

该式子就是凯利公式最常见的形式。在之前举的例子当中，赌客押中的概率 p 为 50%，押错的概率 q 为 50%，赔率 b 为 2，计算得出最优的 f 是 0.25，即 25%，与图 10-2 所示相符。

在搞清楚凯利公式的推导过程之后，其实可以很方便地将基本的凯利公式推广到其他形式。例如，当我们研究一个择时策略的时候，该策略每一次交易都可以看作一次赌局，而针对这些交易需要决策出最优的仓位，也就是赌局中的最佳押注比例。与赌博最大的不同点在于，执行择时策略的过程中，如果由于判断错误而产生损失，也不会一次性损失掉全部资金。对应的赌局中，押金就不再因为押错而全部损失掉。在建模时，研究者需要对原先的设定进行修改。

基于这种情况，仍然假设存在一个赌局，赌客可以一直持续地进行赌博，并且可以在每一次赌局中押注任意的数额。对赌客而言，押中的概率为 p，押错的概率为 q。当赌客押中时，押金退还，赌客额外获得押金的 b 倍；当赌客押错时，押金退还，赌客额外损失押金的 a 倍。⊖

如果赌客采用按固定比例下注的押注手法，每一次的押注金额都是当时金额的 f，f 为某一比例值。当赌局进行了 N 次之后，赌客资金量的期望值可以表示为

$$E(W)=(1+bf)^{pN}(1-af)^{qN}$$

赌客的决策仍然是选择合适的 f，使得 $E(W)$ 最大。最大化 $E(W)$ 实际上与最大化 $\log(E(W))$ 等价，而 $\log(E(W))$ 的形式可以描述为

$$\log(E(W))=pN\log(1+bf)+qN\log(1-af)$$

$\log(E(W))$ 的一阶导数为

$$\frac{\partial\log\left(E\left(W\right)\right)}{\partial f}=\frac{pNb}{1+bf}+\frac{qN\left(-a\right)}{1-af}$$

最高点的 f 取值，应该满足 $\log(E(W))$ 的一阶导数为零，方程如下

⊖ 如果用这样的赌局类比重复性择时策略，那么 a 一般小于 100%，当策略执行涉及杠杆时，也只影响到 f 的大小，而不影响 a。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

$$\frac{pNb}{1+bf}+\frac{qN(-a)}{1-af}=0$$

可以得到

$$f=\frac{p}{a}-\frac{q}{b}$$

具体推导过程如下：

现有

$$\frac{pNb}{1+bf}+\frac{qN(-a)}{1-af}=0$$

则

$$\frac{pb}{1+bf}=\frac{qa}{1-af}$$

$$pb(1-af)=qa(1+bf)$$

$$(qab+pab)f=pb-qa$$

由于

$$p+q=100\%$$ ⊖

则

$$(ab)f=pb-qa$$

$$f=\frac{pb-qa}{ab}$$

即

$$f=\frac{p}{a}=\frac{a}{b}$$

更改一下之前猜硬币例子中的具体设定，赌客押中和押错的概率仍然各为 50%，赌客押 1 元，当赌客押中时，押金退还，赌客额外获得 0.2 元，赌客资金变为 1.2 元；当赌客押错时，押金退还，赌客额外损失 0.1 元，赌客资金变为 0.9 元。对应到参数的设置，p 为 50%，q 为 50%，b 为 0.2，a 为 0.1。与原例子相比，盈亏比保持一致依然是 2，但是变化程度都同时变小了。

代入新的公式中进行计算，最佳押注比例 f 的结果为 2.5。如果该例子中赌客不能借贷资金，那么其能够使用的押注比例仍然是 0 ～ 100%。图 10-3 给出了押注比例与 100 次赌局之后资金量期望值的趋势变化关系，可以看到，当押注比例由 0 增加到 100% 时，交易后的资金量期望值是随之上升的。也就是说，赌客的最优决策是每一次赌局都将全部资金押上。

⊖ p 和 q 都应该为 0 ～ 100%，且满足 $p+q=100\%$

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

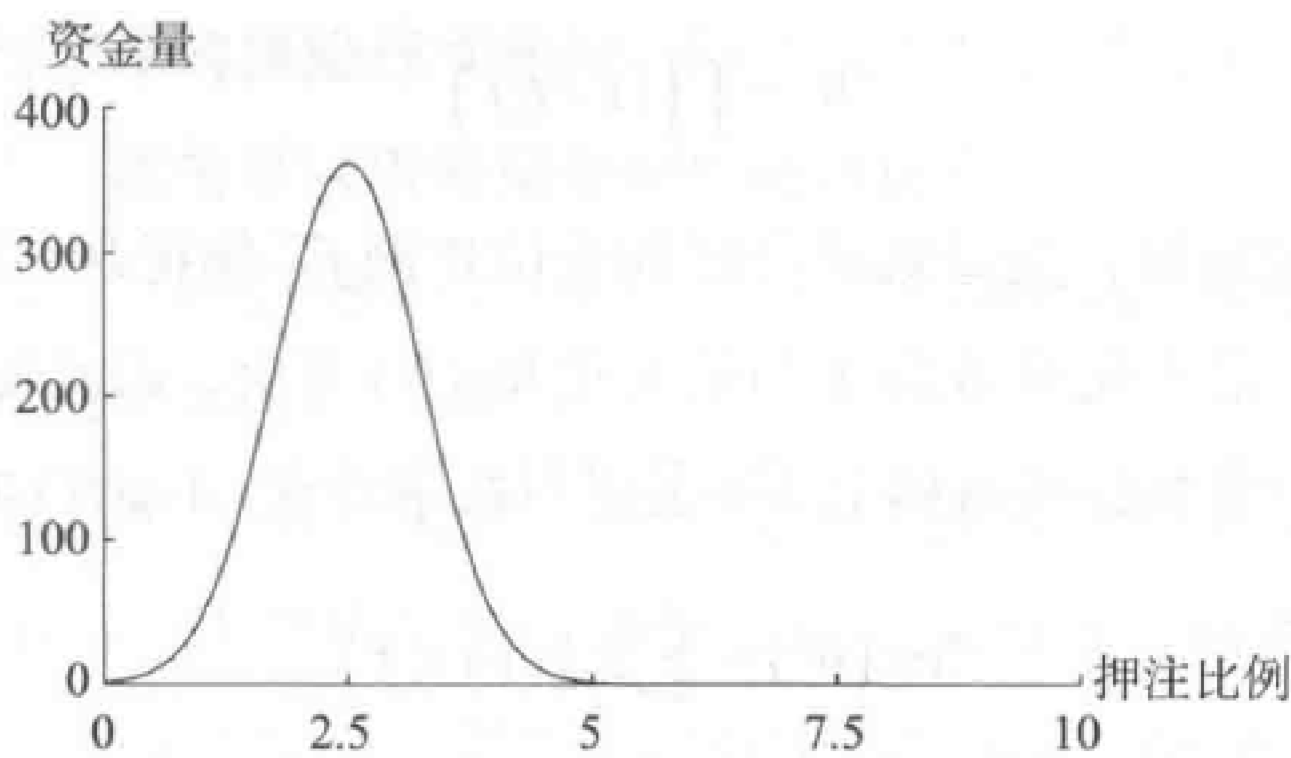


图 10-3 押注比例与资金量的趋势变化关系

当然，这个赌局的例子实际上是为了类比择时策略的情况，押金不会在押错时完全损失掉。那么，为了类比某些交易资产可以使用杠杆的实际情形，这里讨论一下赌客可以借贷的情况。由于依然要满足押注决策不会让资金一次性损失掉的前提，因此 $1-af$ 应该大于零，f 的合理取值范围为 0 ~ $1/a$。在该例子当中，a 取为 0.1，则押注比例不应超过 10，这也正是图 10-3 中横轴范围取为 0 ~ 10 的原因。最佳押注比例 f 的结果为 2.5，也就是使用 2.5 倍的杠杆。

在凯利公式的实际使用中，许多策略研究者、使用者会根据历史数据和回溯测试结果，统计出量化交易策略的盈利概率 p 和损失概率 q，通过求平均值的方法得到策略盈利时的盈利程度 b 和策略损失时的损失程度 a。将这些值代入凯利公式，就可以得到每一次交易的最佳仓位比例。如果结果超出杠杆率限制，则使用最大的仓位。

这样做当然是可以的，但是既然已经知道了凯利公式的内涵以及求导过程，其实我们完全有能力根据实际的量化交易策略回溯测试结果进行更精密的处理。至少可以基于回溯测试中每一次交易的盈亏状况，计算得到历史数据下最优的策略仓位设置。⊖而基于历史与未来一致的假设，这一结果就是可以实际使用的最佳仓位。

假设使用某一个量化择时策略，在历史数据上一共经历了 N 次交易，在不涉及杠杆、满仓交易的情况下，第 i 次交易的收益率为 r_i。则这段样本时期内执行量化交易策略所获得的最终净值为

⊖ 相应的推导过程与拉尔夫·文斯（Ralph Vince）推导 Optimal F 的过程有相同之处，感兴趣的读者可以参阅相关资料。但是作者更倾向于将这些模型都归为凯利公式的大类当中，在作者看来，模型背后的逻辑内涵是认知一个模型的根本所在。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

$$W=\prod_{i=1}^{N}\left(1+r_i f\right)$$

所需要做出的决策，就是选择合适的仓位比例 f，使得最终净值 W 最大。

同样的，由于最大化 W 实际上与最大化 $\log(W)$ 等价，因此先将 W 对数化，从而在保留 W 原有趋势特性的基础上更容易进行数学处理。$\log(W)$ 的形式可以描述为

$$\log\left(W\right)=\sum_{i=1}^{N}\log\left(1+r_i f\right)$$

$\log(W)$ 的一阶导数为

$$\frac{\partial\log\left(W\right)}{\partial f}=\sum_{i=1}^{N}\frac{r_i}{1+r_i f}$$

$\log(W)$ 的二阶导数为

$$\frac{\partial^2\log\left(W\right)}{\partial f^2}=\sum_{i=1}^{N}\frac{-r_i^2}{\left(1+r_i f\right)^2}$$

由于每一个叠加项的分子分母都是平方形式，因此在 N 个收益率 r_i 不全部为零的情况下[⊖]，$\log(W)$ 的二阶导数是严格小于零的。基于这样的数学特性，$\log(W)$ 随 f 变化的形状也应该是先上升后下降，且存在一个总体上的最高点。我们可以拿图 10-3 中的曲线形状作为参考，来假想 $\log(W)$ 的形状与特征。

而这个最高点的 f 取值，也应该满足 $\log(W)$ 的一阶导数为零，即这一点上的切线应该平行于横坐标。通过求解 $\log(W)$ 的一阶导数等于零这一方程，可以得到最优的 f 取值，具体方程如下

$$\sum_{i=1}^{N}\frac{r_i}{1+r_i f}=0$$

不同于之前所举的两个例子，求出这一方程的解析解较为困难。好在 $\log(W)$ 的二阶导数具有比较好的数学特性，在一般情况下都严格小于零。因此我们可以计算出 f 取为零和取为最大值时的一阶导数值，如果符号一致，则最优仓位比例 f 就在边界上；如果符号不一致，则采用二分法等迭代算法得到 f 的最优结果。

当然，采用网格搜索的方法试出最优解的近似值也并不麻烦。可以说，在实践中根据交易策略的回溯测试计算出最优仓位比例是一件相对而言比较容易完成

⊖ 当然，也很难想象会存在一个正常的择时策略，可以使得每一次交易都不赚不亏。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

的工作。在历史与未来的风险程度保持一致的假设下，通过上面所述的模型方法计算出的仓位决策，就是可以用来指导交易的最优仓位了。

将凯利公式及其应用叙述至此，应该可以很明确地认清凯利公式的核心内涵了。作者将其大致归纳为两点：

（1）凯利公式试图规避的风险，是一次交易就将可用资金全部损失掉的风险。在赌客确保不会一次性出局的基础上，只要赌局对于赌客而言在概率上有利，赌客就具有将资金增值的潜在能力。而凯利公式通过固定比例的押注方式，实现了不会一次性出局的目标，固定比例是该公式的一个核心设置。

（2）凯利公式寻求的目标是，经过多次赌局之后，实现赌客资金量的最大化，即赌客收益的最大化。在推导过程中，最大化往往针对的是收益的期望，这是一个更数学化也更容易处理的做法。

作者在本节的内容中，不厌其烦地改变条件，推导凯利公式的具体应用形式，其实也正是希望告诉读者，当我们从内涵上理解了凯利公式的核心所在之后，完全可以根据量化交易策略的研发环境，自主地推导公式，选取最为合适的模型构造方法。其他形式如分布条件下的凯利公式⊖等，不再一一赘述，有兴趣的读者可以自行推导或寻找资料进行了解。

可惜的是，某些量化交易策略的研究者在实际工作中只是简单地套用凯利公式，甚至错误地使用了凯利公式的基本形式，这往往是对该公式的推导过程不求甚解的结果。而实际上，推导凯利公式涉及的这些数学模型其实并不复杂，掌握整个模型也不是一件非常困难的事情。

因此，作者在此建议读者，在量化交易策略的研发过程中，不管是使用凯利公式，还是处理更为复杂的数学模型和交易策略，都应该更多地关注策略模型内在的含义，由内及外地加以应用。而不是简单地套用模型或公式，丧失对其的控制能力和改进能力。

最后，讨论一下凯利公式相关的知识点。虽然在之前的假设环境下，当赌局在概率上对赌客有利时，凯利公式是一个非常优秀的仓位决策方法。但是在实际操作中，当资金量下降到一定程度时，即使没有完全为零，也会失去继续玩下去的资格。例如赌客从 1 万元开始赌博，输到 1 元钱时，庄家可能就不会再接受 1

⊖ 欧内斯特·陈（Ernest Chan）的两本书对分布条件下的凯利公式均有介绍。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

元的押注了。原例子中允许所有小数金额，如0.000 001元的假设，并不是一个合乎实际的假设。

类似的情况还有倍赌[1]押注方式，简单情形如一个赔率为1的赌局[2]中，赌客用1元下注，输了就接着押注2元，再输了就接着押注4元，循环往复。就算前面全部押错，只要有一局押对了，整体来看就可以盈利1元。而从概率上讲，赢一局这一事件是必然会发生的。因此粗看起来，这一押注方法几乎是必胜之法。但是关键在于，若干次连续押错之后，赌客可能就没有足够的钱进行加倍了。

当形势不利时，凯利公式的赌注越来越小，倍赌的赌注则越来越大，两者有明显的分野。然而在现实中，既存在使用凯利公式这一类设置的量化交易策略，也存在使用倍赌类设置的量化交易策略。如何根据实际情况进行选择，也是需要建立在掌握模型公式、理解其优点和不足之处的基础上的。

另外，在本节的几个例子中，赌局都在概率上对赌客有利，量化交易策略的回溯测试也应该取得整体正收益。而如果赌局在概率上对赌客不利，或者是量化交易策略回溯测试的整体收益为负时，那么套用相同公式算出的f应该为负数。针对赌局，可以理解为赌客无法在概率意义上获利。针对量化交易策略，可以理解为，使用方向相反的仓位，就可以取得最大正收益。

10.2 实用的仓位决策方法

上一节所阐述的凯利公式正是量化择时策略进行仓位决策的一种行之有效的方法。在通过上一节内容所形成的基本认识的基础上，我们在这一节中进一步说明实际研发中的仓位决策是如何做出的。

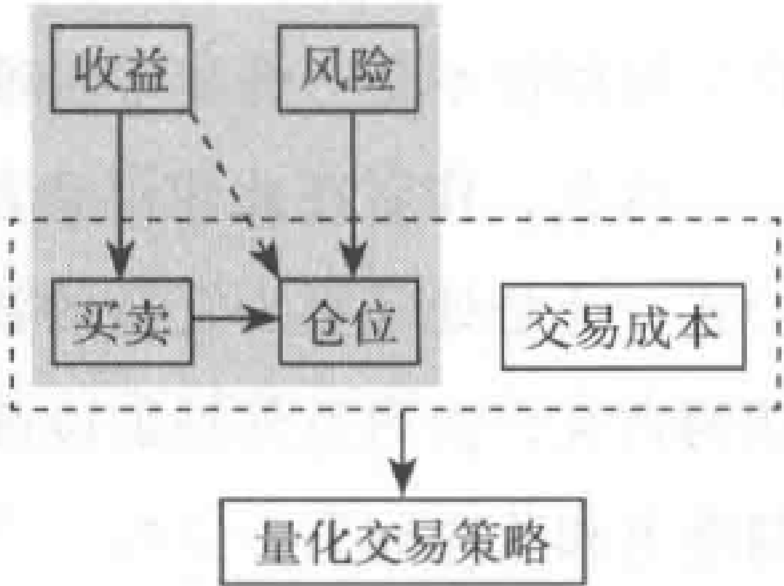


图 10-4 仓位决策在量化择时策略研发流程中的位置

图10-4通过灰色区域说明了仓位决策在量化择时策略研发流程中的位置。可以看到，仓位的设定既依赖于买卖决策，如第4章、第7章涉及的择时策略中的决策行为，也依赖于量化交易策略对风险的控制。

[1] 英文为Martingale，在一些环境下翻译为“鞅”，这里应该取“倍赌”之意。
[2] 这是一个对赌客和庄家在概率上十分公平的赌局。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

需要注意的是，在凯利公式的设计中，控制的风险是一次性损失掉所有资金的可能性，这是一个相对而言程度比较弱的风险控制行为。对该公式而言，仓位优化更重要的目标在于最大化交易策略的整体收益。图 10-4 灰色区域中由收益到仓位的虚线箭头，其含义就在于此。

因此，在凯利公式的逻辑下，风险控制对仓位的影响是偏小的。更重要的是，由于凯利公式控制的风险是一次性损失掉所有资金的可能性，因此当实际的风险大于预估的风险时，凯利公式得到的仓位设置可能就会导致资金一次性损失掉的情况出现。而由于量化交易策略的研究当中不可避免地存在过度拟合的情况，因此回溯测试下策略所表现出的风险往往都小于策略在未来所具有的风险。在这种情况下，凯利公式的风险控制能力是值得怀疑的，单纯地使用凯利公式本身就具有很高的风险。

在实际的量化交易策略研发过程中，虽然凯利公式具有很好的内在逻辑，但是其带来的“安全感”过低，所以我们很难仅仅使用凯利公式来确定策略的最优仓位。解决这一问题最为直接的途径就是提高风险控制的程度，使其不仅仅满足于规避一次性损失所有资金的风险。具体的做法是直接使用风险度量指标来描述风险水平，通过调整仓位设置，保证量化交易策略的风险不超过某一具体水平。

关于风险度量指标，本书已经在第 9 章的 9.1 节和 9.2 节中进行了一定的介绍，这里不再赘述。对一个完整的量化交易策略⊖而言，在买卖设置相同的前提下，设定不同的仓位决策，就会具有不同的风险水平，具体表现为不同的风险度量指标结果。那么要做的，就是确定合适的仓位设置，使得风险度量指标的结果等于研究者事前设定的数值。

当然，在使用风险度量指标确定仓位决策时，一样存在着回溯测试下策略所表现出的风险小于策略在未来真实具有的风险的问题。因此，使用某个风险度量值作为阈值求出的仓位决策，往往也会导致超过该阈值的实际风险。在实际中可以通过调低风险度量指标的阈值取值来平衡。

在图 10-4 的研发框架中，不论是使用凯利公式，还是使用纯粹的风险度量指标来确定仓位决策，都是建立在买卖决策已经建立的前提下的。我们知道，最优的买卖决策是基于收益最大化导出的。因此，即使是使用纯粹的风险度量指标

⊖ 指的是同时做出买卖、仓位决策的量化交易策略。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

来进行仓位决策，整体而言也同时考虑了收益与风险。

而凯利公式与风险度量指标的差异点在于，凯利公式导出的仓位决策更看重收益，风险度量指标导出的仓位决策更看重风险。两者的侧重点不同，是一组比较互补的仓位决策方式。在实际工作中可以同时考虑这两种方式，综合两个仓位决策结果来决定最后使用的仓位决策，如图 10-5 所示。

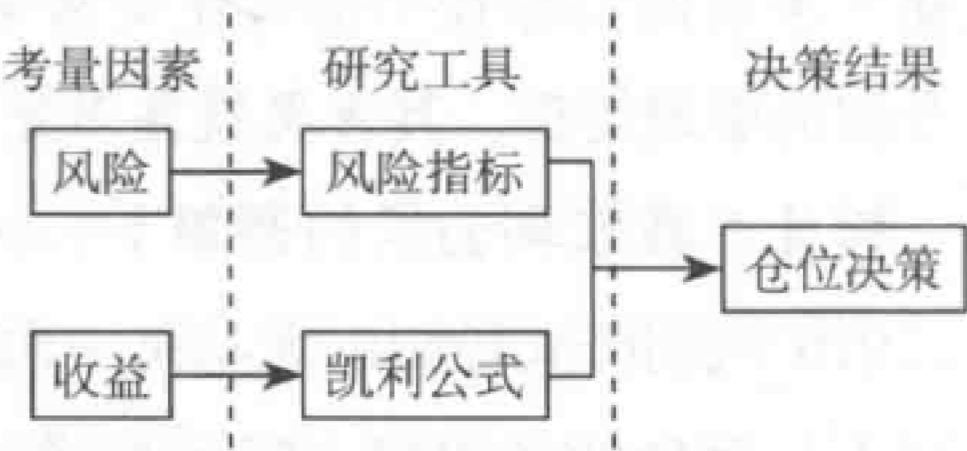


图 10-5 仓位决策的构建

只是在实际的量化交易策略研发中，使用凯利公式所得到的最佳仓位往往都相对较重，通过风险度量指标限制所导出的仓位往往会偏小一些。只要量化交易策略的使用者不是太过激进，一般都会更多地考虑较小的仓位，或至少给予小仓位更多的考量权重。

当然，凯利公式与风险度量指标这两种方法得到的仓位决策孰大孰小并不是绝对的。鉴于两种仓位决策手段都不难实现，稳妥的办法是计算两个标准下的仓位决策结果，综合进行考虑。比较简单的做法就是取更小的那个仓位设置。

$$f=\min\{f_{\text{凯利}}, f_{\text{风险}}\}$$

式中，$f_{\text{凯利}}$是使用凯利公式得到的仓位设置；$f_{\text{风险}}$是使用风险度量指标限定得到的仓位设置。$f_{\text{凯利}}$和$f_{\text{风险}}$都代表整体资金量的一定比例，因此仓位设置结果f也是一个比例。

上文提到了，不管是凯利公式还是风险度量指标，这两种方法都存在实际风险大于预估风险的问题。表现在现实中，就是回溯测试下策略所表现出的风险，往往都低于策略实际的内在风险，也就低于策略在未来实际表现出的风险，因此将其作为仓位决策的标准时，得到的仓位往往都是过高的。

解决的办法，一是如同前面说的，通过调低风险度量指标的阈值取值来得到一个相对更低的仓位结果，也就是主动调低回溯测试下的风险承受能力；二是干脆直接将仓位设置结果f按比例减小。一般而言，第二种方法虽然直接，而且对凯利公式和风险度量指标两种方法同时有效，但是经过一个中间步骤之后，其内在逻辑性减弱了。当然，第一种方法中如何调低风险水平也是一个缺乏定量方案的问题，研究人员如何选择还需要自行摸索。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

第 11 章

仓位优化的择时策略

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

11.1 仓位优化的均线趋势策略

上一章介绍了量化择时策略进行仓位决策的具体模型和方法。在本章中，作者将基于第 7 章所介绍的推进的择时策略，进一步构造合理的仓位决策，将买卖和仓位两者相结合，使之成为更加完善的量化交易策略。

首先论述的是第 7 章 7.1 节所述的均线趋势策略。推进分析的框架依然保留，额外要做的就是在推进分析下加入仓位决策的部分。

双均线趋势策略的具体定义依然和第 4 章 4.2 节中的保持一致。在交易决策的时点需要根据已知数据计算短期均线和长期均线两个值。当短期均线高于长期均线时，判断交易决策时刻的趋势为上涨趋势，认为后市会继续上涨，因此看多。而当短期均线低于长期均线时，判断交易决策时刻的趋势为下跌趋势，后市继续下跌，因此看空。具体可参见图 4-5 和相关叙述。

数据方面也和前面所选取的数据一致，交易资产是以中证指数公司编制的沪深 300 指数为标的物的沪深 300 股指期货主力合约，收集得到从 2010 年 4 月 19 日到 2015 年 6 月 30 日的日收盘价数据。

依照样本时间段内的期货市场情况，将股指期货的交易保证金设为 15%。由于样本期到 2015 年 6 月 30 日结束，此时股指期货非套期保值持仓交易的新规还未发布，因此将保证金设为 15% 是一个与实际情况出入不大的处理。

大的推进分析框架也仍然维持了第 7 章 7.1 节中的设定。图 11-1 展示了均线趋势策略的推进分析在这段时间的样本下是如何进行的。2010 年 4 月 19 日到 2010 年 12 月 30 日的日收盘价数据作为备用数据用来计算短期均线和长期均线，数据长度为 172 日，也就是说只要长期均线的长度在 173 日以下，备用数据都可以支持整个推进分析的计算顺利完成。需要注意的是，这部分数据的使用在图 7-1 中没有具体显示出来。

与第 7 章 7.1 节中研发过程不同的地方在于，在每一行的白色框优化期计算中，不仅要计算出最优的均线组合，还要基于该均线组合所形成的买卖决策，计算得到相应的最优仓位设置，并将均线组合设置和仓位设置同时应用于对应的灰色框上。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

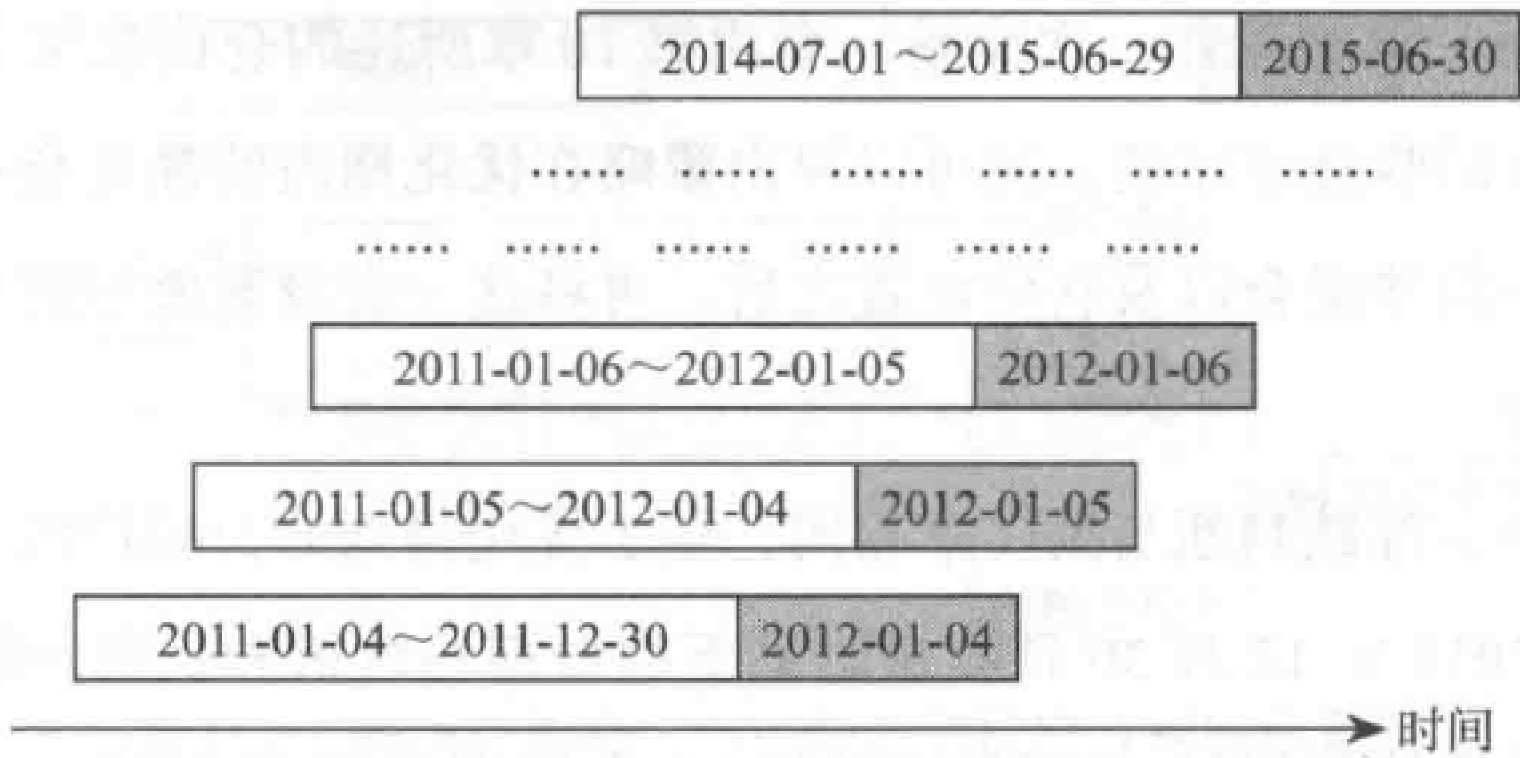


图 11-1 仓位优化的均线趋势策略的推进分析流程

基于这样的目的，我们在白色框的优化过程中，不再依照习惯将一次趋势变化视作一次交易，每一次由涨变跌或者由跌变涨时进行一定数额的买卖操作，而是将每一日都看作一次交易，在每一日都变更具体的仓位设置。

这一设定的主要原因在于，推进分析中灰色框的步长为一日，也就是检验期仓位变更的频率是一日一次，为了保证优化与检验两个步骤在仓位逻辑上的一致性，就依此设定优化期的仓位变更也是一日一次。

以第一行操作为例，第 7 章只是使用网格搜索优化出了盈利能力最好的长短均线组合，就直接将其用来计算 2011 年 12 月 30 日收盘时刻下的短期均线值和长期均线值。当短期均线值大于长期均线值时，就对 2011 年 12 月 30 日收盘到 2012 年 1 月 4 日收盘的走势看多；当短期均线值小于等于长期均线值时，就对 2011 年 12 月 30 日收盘到 2012 年 1 月 4 日收盘的走势看空。买卖交易的杠杆率设为 100% 不变。

而在本节中，由于在均线趋势策略中增加了仓位决策的部分，因此在优化搜索出盈利能力最好的长短均线组合之后，还需要再一次使用白色框的数据，优化出相应的仓位设置。

具体而言，针对 2010 年 12 月 31 日到 2011 年 12 月 30 日的日收盘数据，首先通过搜索不同均线组合的交易情况来寻找最优的均线趋势策略参数，这与第 7 章的做法是一样的。然后，将 2010 年 12 月 31 日到 2011 年 12 月 30 日设想为策略交易期，得到最优均线组合下每一日的持仓方向。这样，在白色框的时间段上，可以获得 244 个交易日的日度收益率。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

我们将每一日都看作一次交易，套用第 10 章所述的仓位决策方法，基于这个长度为 244 的收益率序列，就可以得出策略在优化期内的最优仓位设置了。得到最优的长短均线组合以及仓位设置之后，再将这一套设置应用到对应的灰色框中，进行检验。

灰色框中，涨跌判断与第 7 章相同，基于优化得到的长短均线组合，直接将其用来计算 2011 年 12 月 30 日收盘时刻下的短期均线值和长期均线值。当短期均线值大于长期均线值时，就对 2011 年 12 月 30 日收盘到 2012 年 1 月 4 日收盘的走势看涨；当短期均线值小于等于长期均线值时，就对 2011 年 12 月 30 日收盘到 2012 年 1 月 4 日收盘的走势看跌。

2011 年 12 月 30 日收盘到 2012 年 1 月 4 日收盘期间的仓位，则直接使用前面优化得到的仓位比例设置，看涨时仓位为多，看跌时仓位为空。也就是说，仓位的方向和仓位的比例，共同构成了整体的仓位决策结果。

在灰色框的建仓时点，也就是 2011 年 12 月 30 日收盘时，根据上一个交易日的仓位情况，计算其与目前的仓位决策的差值，就是当前所需要进行的具体交易数额。例如，检验期从 2011 年 12 月 30 日收盘时开始，那么之前的仓位为 0，如果计算得到的仓位按比例计算为多 10 手，那么对应的交易操作就是买入 10 手。假设之前的仓位为 9 手，那交易操作则应该是买入 1 手，依此类推。

具体说明了推进分析第一行包括怎样的流程之后，我们再回到图 11-1 所示的推进分析整体流程中。在进行完一行的操作之后，将白色框和灰色框都分别向后推进一日，重复这种样本外检验的过程，直到最后一行为止。每一个灰色框也就是每一个交易日之间，都存在着重新计算仓位决策、进行买卖操作从而满足新的仓位要求的过程。图 11-2 给出了推进分析中每一行的流程示意，以供参考。

最后一行中被检验的灰色框为 2015 年 6 月 29 日收盘到 2015 年 6 月 30 日收盘，也就是触及样本数据的最后一个数据。总的来看，在推进分析的过程中，模拟买卖的行为从 2011 年 12 月 30 日收盘时开始，到 2015 年 6 月 30 日收盘时结束，一共交易三年半的时间，也就是所有灰色框所覆盖的区间。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

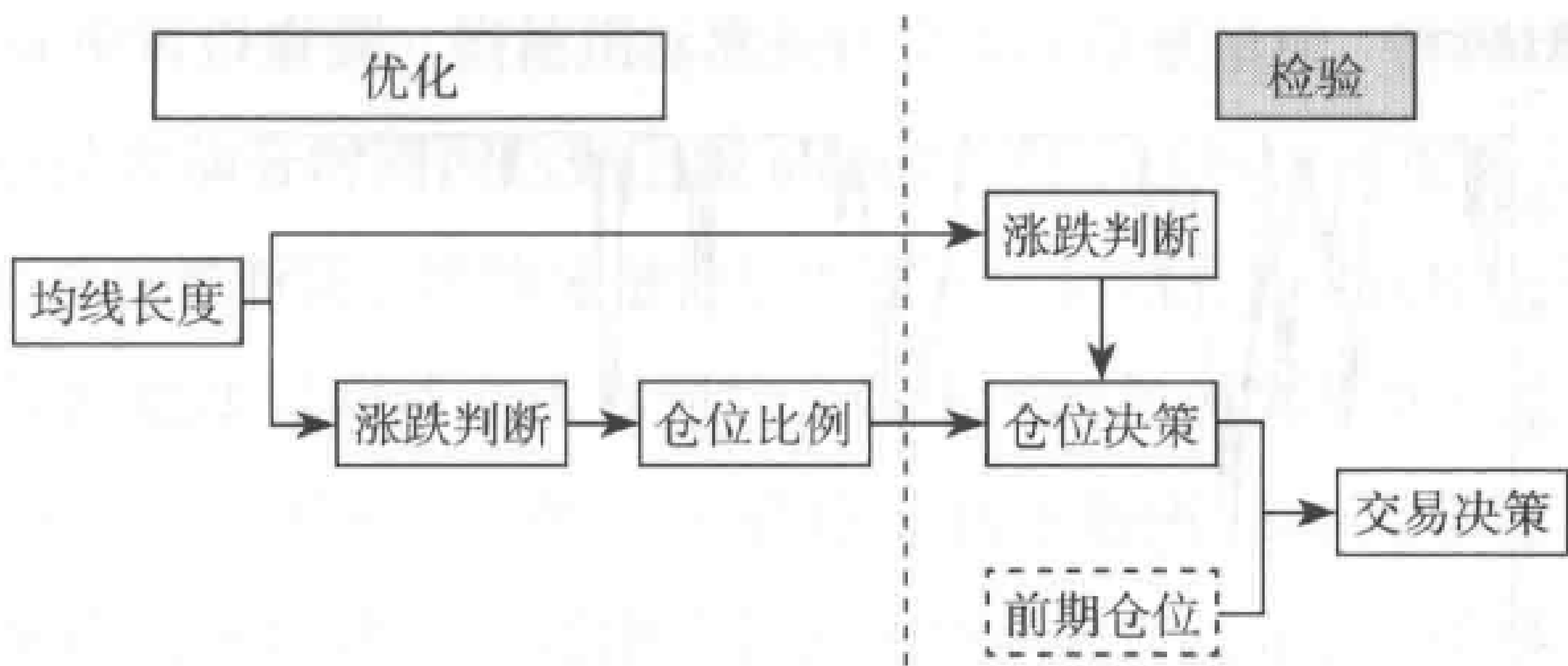


图 11-2 推进分析每一行中优化和检验过程的示意图

之前曾提到，在决定了最优的长短均线长度之后，不同的风险控制程度会导致不同的仓位设置结果。这里我们使用第 10 章 10.2 节介绍过的仓位决策方法，通过使用凯利公式与风险度量指标这两种方法得到各自的仓位设置结果，再取较小的那个结果作为最后的仓位决策。

$$f=\min\{f_{凯利}, f_{风险}\}$$

在凯利公式的设计中，其所控制的风险是一次性损失掉所有资金的可能性，这是一个相对较为固定的风险控制水平。由于涨跌判断已经固定，收益率序列也就固定了，因此最大化收益的目标也会使得凯利公式的仓位结果相对固定。在这种情况下，导致仓位结果不同的主要原因就是风险度量指标的设置及其阈值的选取。

首先对凯利公式所导出的仓位结果做统一观察。图 11-3 描述了基于推进分析中每一行的白色框数据，分别使用凯利公式计算得到的最优的仓位比例设置，即 $f_{凯利}$。推进分析一共进行了 845 行，图 11-3 所展示的也就是一个长度为 845 的时间序列。而前文描述数据时已经说明了，股指期货的交易保证金设为 15%，因此仓位比例的可选择范围就是 0 ～ 6.6667。⊖

可以看到，大部分情况下，凯利公式计算得出的仓位比例都达到了上限 6.6667。在部分时期，仓位结果经历了快速下降然后又快速回升的过程，表现在图形中就是一个个的 V 字形态。凯利公式下的仓位比例在 2012 年 10 月 12 日达到最低，为 3.0379。虽然是全局的低点，但是这一仓位设置为 0 ～ 6.6667 的取值范围仍然不低。

⊖ 不考虑涨跌方向。如果仓位设置中包含方向信息，则取值应为 −6.6667 ～ 6.6667。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

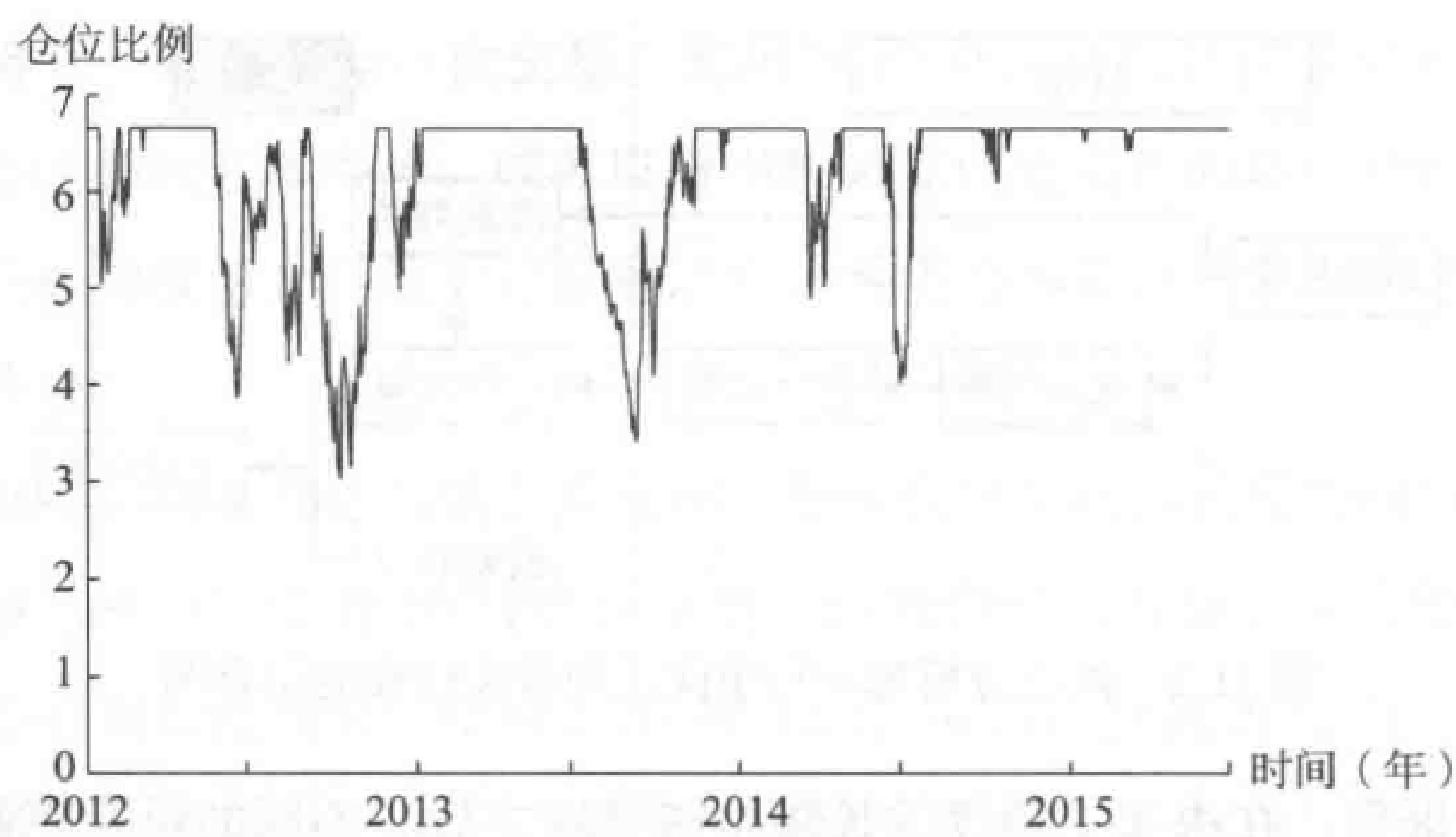


图 11-3 凯利公式下的仓位设置

然后，我们进一步加入风险度量指标的限定。为了所举的案例简便起见，将风险度量指标设定为标准差，阈值水平则分别设为三个不同的数值，0.08、0.04和 0.02。在后面的内容中，作者将分别讨论这三种风险控制水平下均线趋势策略的仓位决策，以及相应的推进分析结果。

11.1.1 标准差阈值 0.08

当标准差的阈值取为 0.08 时，基于推进分析中每一行的白色框数据，分别计算出使得收益率序列标准差等于 0.08 的仓位比例设置，即 $f_{风险}$。这也是一个长度为 845 的时间序列，如图 11-4 所示。

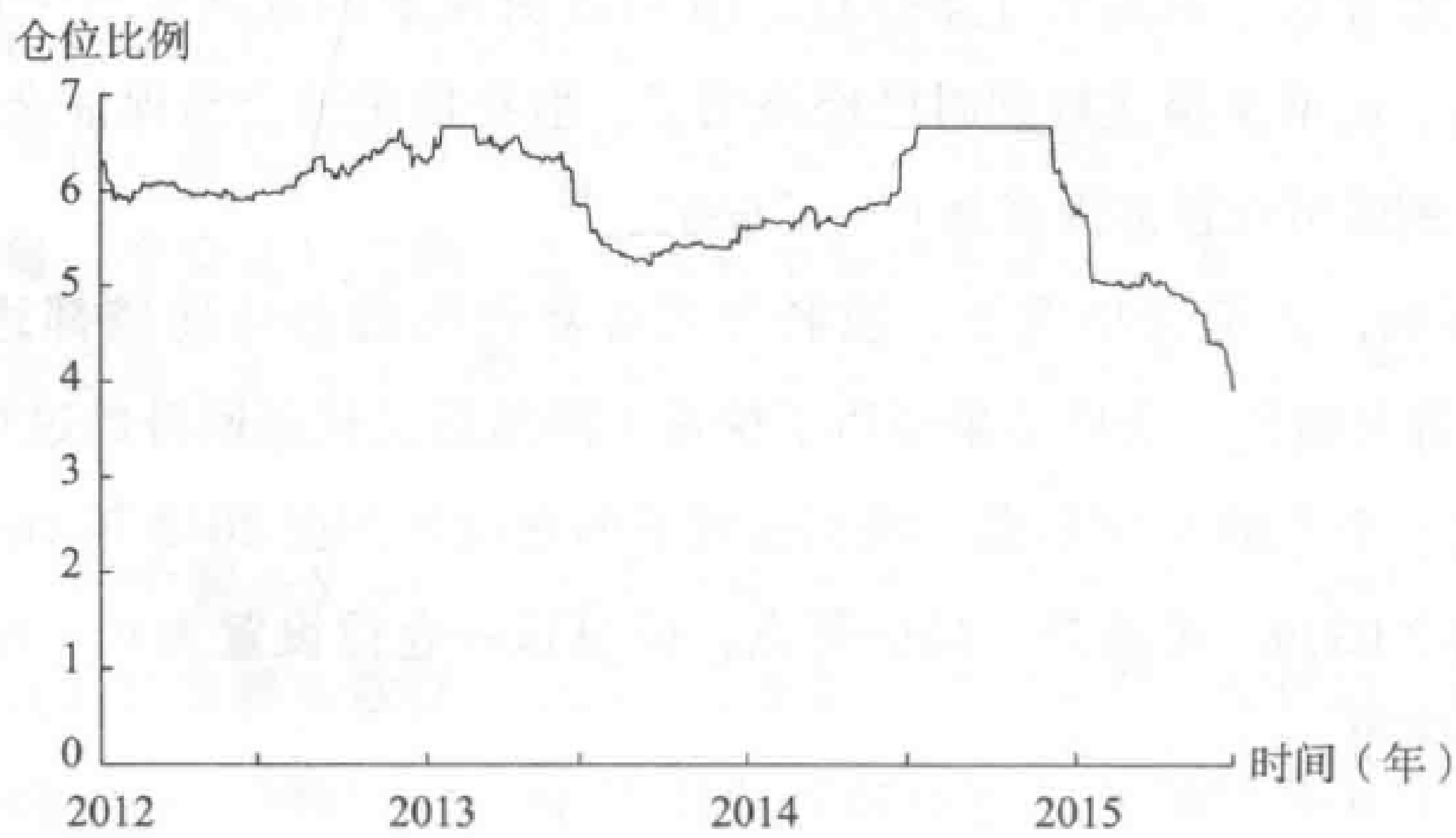


图 11-4 风险度量指标控制下的仓位设置

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

从图 11-4 中可以看到，当使用标准差作为风险度量指标，并限定阈值为 0.08 时，合适的仓位在部分时间内达到上限 6.6667，只不过所占比重不大。相较于凯利公式下的仓位设置而言，风险度量指标控制下的仓位设置变化较为平缓。

接着，我们把使用两种方法得到的仓位比例时间序列放在一张图中加以对比，如图 11-5 所示。实际上，图 11-5 就是图 11-3 和图 11-4 的重叠。可以发现，当标准差的阈值设为 0.08 时，使用凯利公式和风险度量指标分别得出的仓位设置的重合程度是相对较高的。如果使用求较小值的方法得到最终结果，那么仓位决策有一部分来自凯利公式，另一部分来自风险度量指标限制，这是一个看似合理的情况。

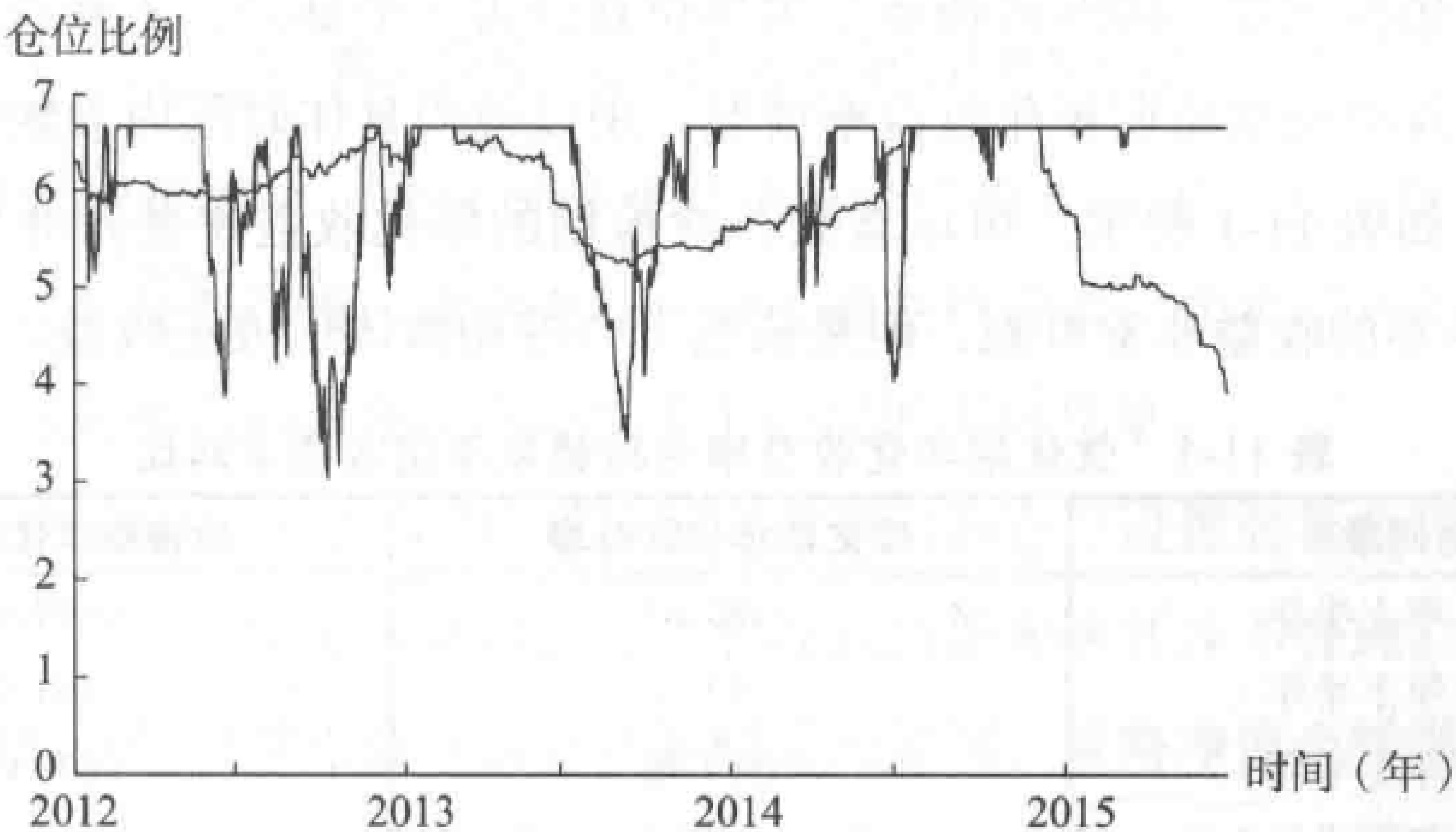


图 11-5 凯利公式与风险度量指标分别得到的仓位设置

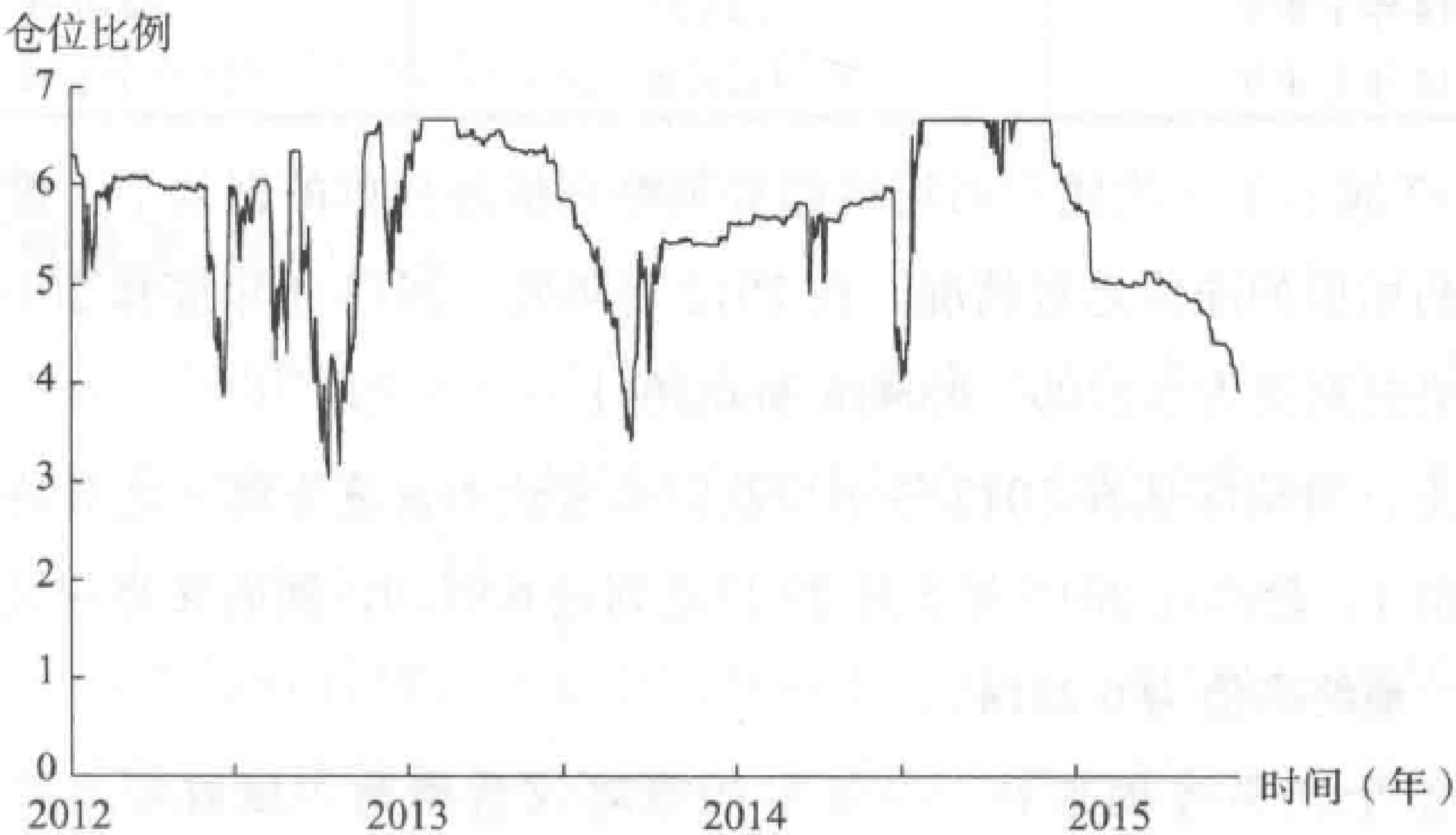


图 11-6 最终确定的仓位设置

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

图 11-6 是逐日求出凯利公式和风险度量指标两种方法计算得到的仓位比例后，取较小值得到的最终仓位比例序列。实际上，图 11-6 也就是图 11-5 中两个曲线图形的下界。图 11-6 所示的仓位比例设置，既是推进分析中白色框逐行优化的结果，也是在灰色框中逐日使用的仓位决策。

设置好仓位之后，先分别查看白色框和灰色框内均线趋势策略的收益表现。将整体综合起来进行计算得知，加入对仓位的决策之后，用来优化的白色框内策略平均的年化收益率为 266.76%，而用来检验策略的灰色框的年化收益率则为 −35.86%。两者差距很大，存在严重的过度拟合现象，甚至导致检验期的收益情况为负。

再将灰色框三年半的时间按半年为单位划分为 7 个部分，考察每半年中的优化期年化收益率和检验期年化收益率情况，用以说明具体时间内参数的过度拟合情况。结果如表 11-1 所示。可以看到，检验期的年化收益率差距非常大，虽然 2014 年下半年的收益非常可观，但是依然无力带动整体形成正收益。

表 11-1　优化期年化收益率与检验期年化收益率对比　（%）

时间段	优化期年化收益率	检验期年化收益率
2012 年上半年	187.63	−97.58
2012 年下半年	91.18	101.94
2013 年上半年	374.70	130.08
2013 年下半年	163.13	−98.46
2014 年上半年	237.46	−43.56
2014 年下半年	214.25	9439.78
2015 年上半年	1125.54	−51.99

图 11-7 展示了仓位优化的均线趋势策略在推进分析的框架下，模拟交易过程以 1 为初始值的净值走势情况。在 2012 年年底、2013 年年底和 2014 年年底，策略的净值分别变为 0.2209、0.0416 和 0.3051。

实际上，策略净值在 2012 年年初就以非常快的速度下降，之后再也没有高于初始净值 1，最高在 2015 年 5 月 27 日达到过 0.9520，随后又是一波速度非常快的下降，最终净值为 0.2114。

基于整个交易过程来看，三年半的策略收益率为 −78.86%，年化收益率为 −35.86%，亏损较为严重。参考第 7 章 7.1 节的结果，即使完全不进行仓位决

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

策，始终保持 100% 的仓位比例，策略同期的年化收益率也有 5.53%。

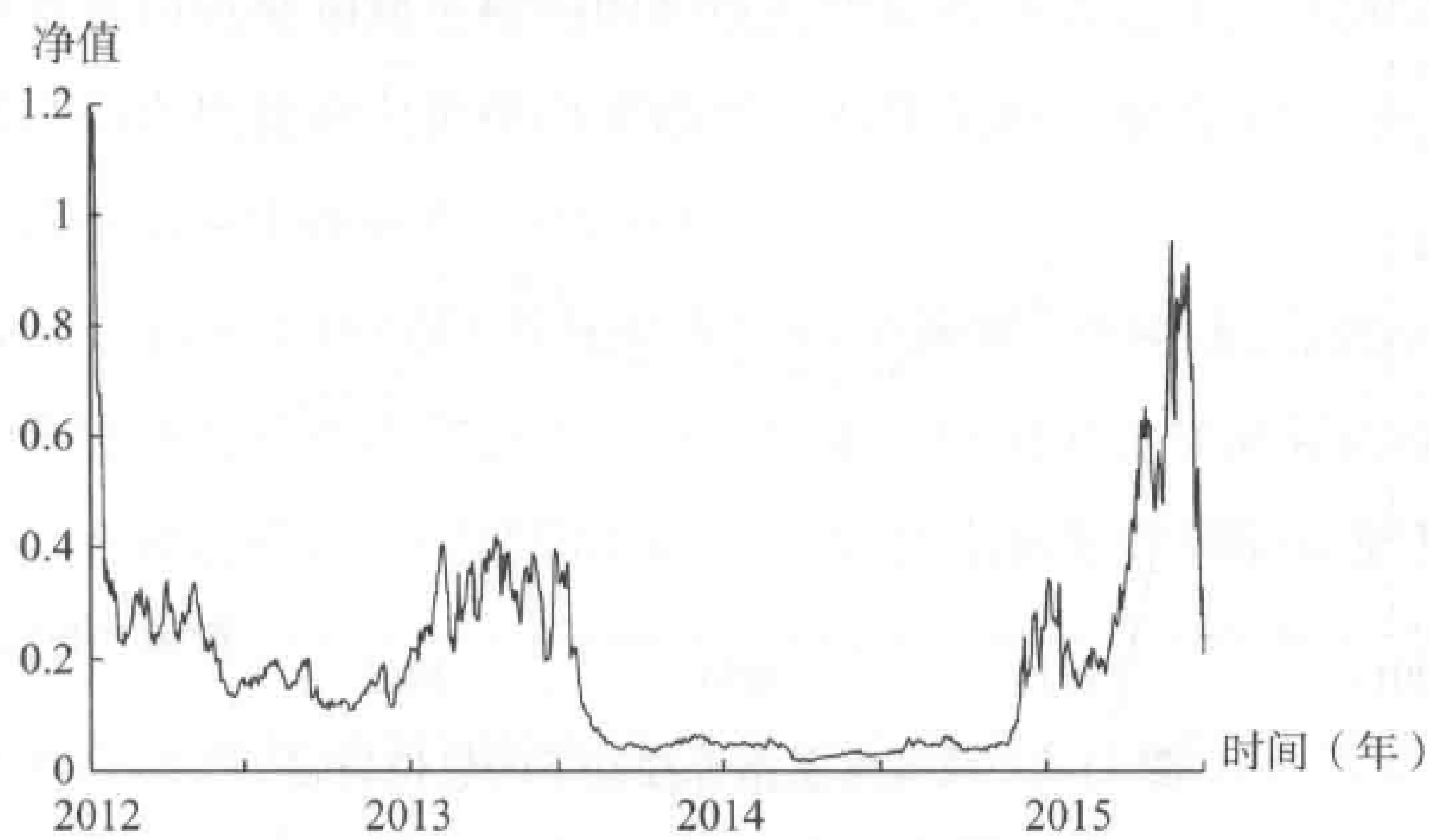


图 11-7　推进分析下仓位优化的均线趋势策略净值走势

原因其实也很明显，就是当标准差阈值设置为 0.08 时，使用既定仓位决策方法得到的仓位设置过重。由白色框中适当的仓位设置，变为了灰色框中过重的仓位设置，这是典型的体现在仓位决策上的过度拟合现象。

虽然使用凯利公式和风险度量指标分别得出的仓位设置的重合程度较高是一个看似合理的情况，但是很显然，最终的收益结果表明其并不合理。纯粹使用凯利公式得到的仓位设置也会导致整体策略无法盈利，策略净值走势类似，这里不再多做展示。这些结果都验证了作者在第 2 章 2.1 节中曾经叙述的观点：在实际使用中，凯利公式所导出的仓位设定往往过于偏激，超过正常风险控制下的最高仓位值，因此仓位仍然与风险的关系更为紧密。

11.1.2　标准差阈值 0.04

当标准差的阈值取为 0.04 时，我们基于推进分析中每一行的白色框数据，分别计算出使得收益率序列标准差等于 0.04 的仓位比例设置 $f_{风险}$，并在图 11-8 中展示这一长度为 845 的时间序列。

从图 11-8 中可以看到，当使用标准差作为风险度量指标，并限定阈值为 0.04 时，合适的仓位比例在 3 左右浮动，整个检验期内都没有达到 6.6667 的上限。这与 0.08 标准差阈值下的情况存在区别。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

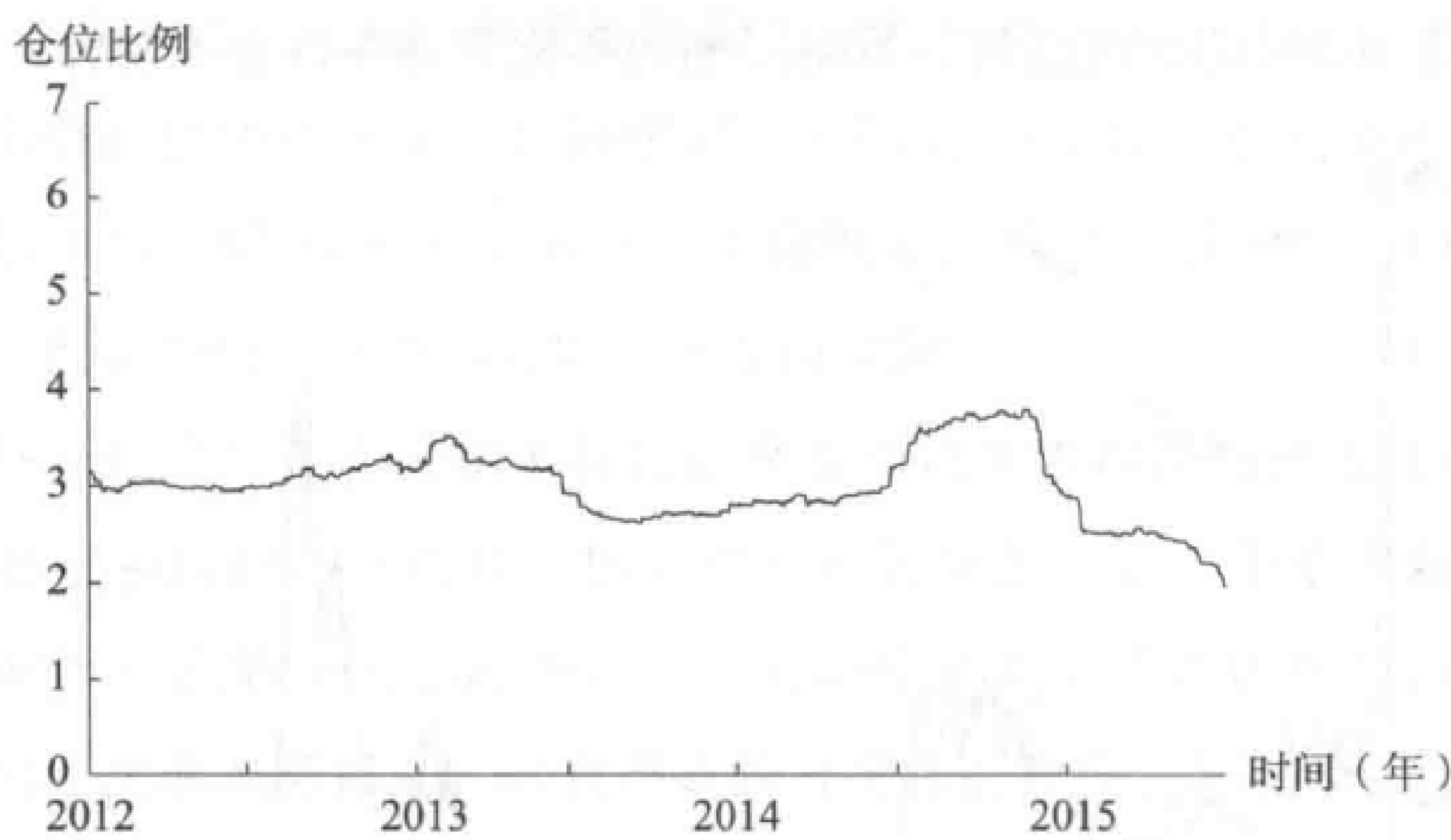


图 11-8　风险度量指标控制下的仓位设置

接着，我们把使用凯利公式与风险度量指标得到的仓位比例时间序列放在一张图中加以对比，实际上就是将图 11-3 和图 11-8 重叠，如图 11-9 所示。很明显，使用这两种方法分别得出的仓位设置几乎不再重合。除一天以外，在其余所有时间里，通过限定风险度量指标数值得到的仓位设置都要小于凯利公式计算得到的仓位设置。

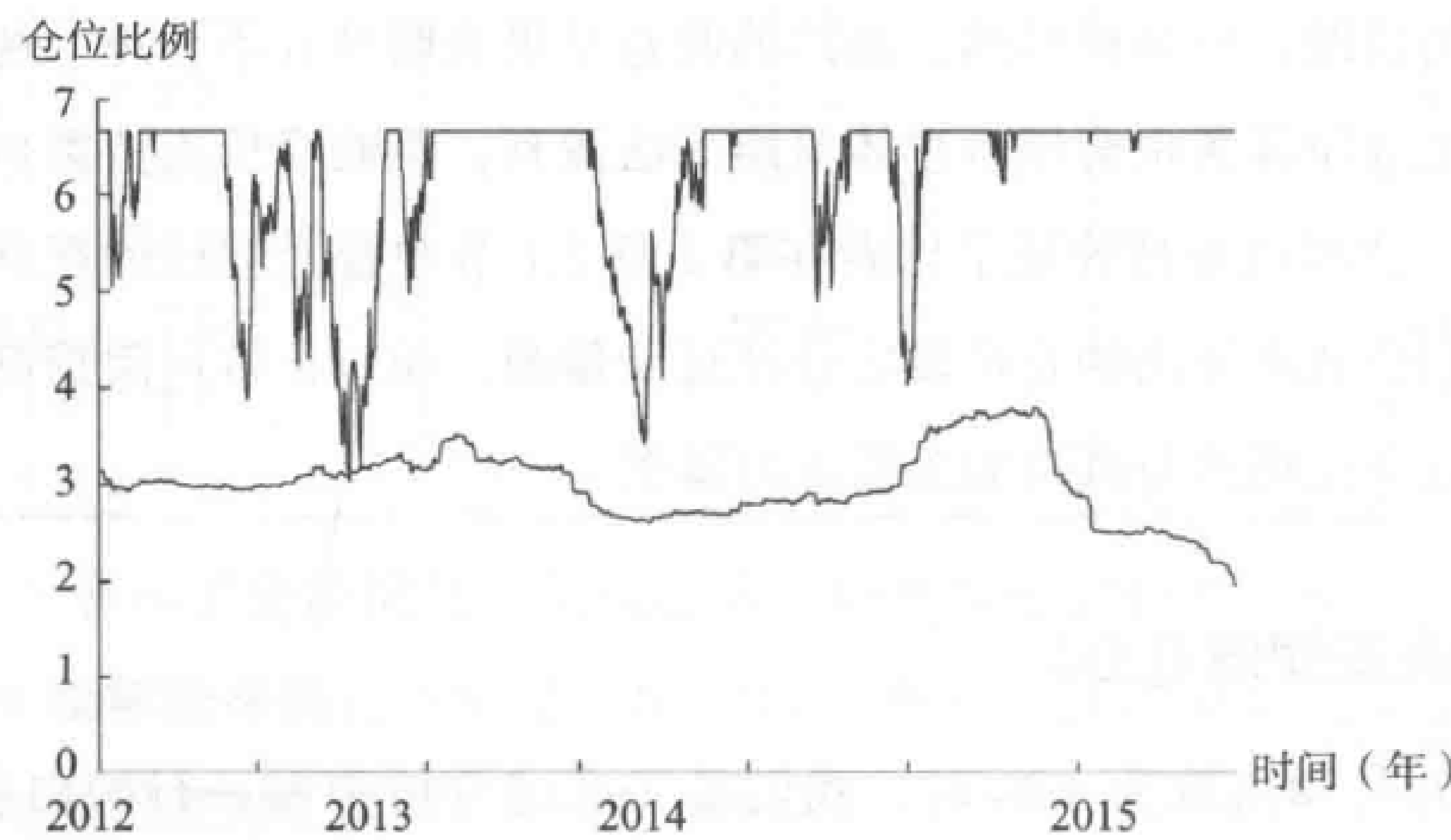


图 11-9　凯利公式与风险度量指标分别得到的仓位设置

那么，当我们使用求取两种仓位决策较小值的方法得到最终结果时，最后使用的仓位设置实际上基本就是基于风险度量指标限制所得到的仓位设置，也就是图 11-8 所示的曲线。由于图形近乎一致，这里就不再多做展示，请读者直接参考图 11-8。该仓位比例设置，既是推进分析中白色框逐行优化的结果，也是在灰

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

色框中逐日使用的仓位决策。

先分别查看白色框和灰色框内的均线趋势策略收益表现。将整体综合起来进行计算得知，用来优化的白色框内策略平均的年化收益率为 132.16%，而用来检验策略的灰色框的年化收益率为 10.53%。

可以看到，相较于标准差阈值取为 0.08 的情况，白色框内的平均年化收益率更小，但灰色框内的年化收益率已由负转正，说明由仓位决策造成的过度拟合程度下降了。这种过度拟合程度的下降，表面上是由改变风险标准导致的，归根到底是缩减仓位所致。

再将灰色框三年半的时间按半年为单位划分为 7 个部分，考察每半年中的优化期年化收益率和检验期年化收益率情况，结果如表 11-2 所示。与表 11-1 对比可知，通过调整标准差的阈值水平，虽然几个原先录得正收益的时期特别是 2014 年下半年，收益水平有所下降，但是原先收益为负的时期的损失情况均有所好转，2015 年上半年甚至由负转正。

表 11-2 优化期年化收益率与检验期年化收益率对比 (%)

时间段	优化期年化收益率	检验期年化收益率
2012 年上半年	103.48	−79.99
2012 年下半年	66.79	98.75
2013 年上半年	162.64	98.31
2013 年下半年	96.91	−85.60
2014 年上半年	120.36	−11.58
2014 年下半年	118.85	1561.51
2015 年上半年	329.40	20.74

图 11-10 展示了仓位优化的均线趋势策略在推进分析的框架下，模拟交易过程以 1 为初始值的净值走势情况。在 2012 年年底、2013 年年底和 2014 年年底，策略的净值分别变为 0.6306、0.3370 和 1.2918。

实际上，在 2012 ～ 2014 年三年间，策略净值水平在大部分时间内都是小于初始净值 1 的。净值从 2014 年末开始呈现出一个上升态势，在 2015 年 6 月 12 日达到最高点 2.7083，随后的一小段时间有一个较为明显的下跌，到 2015 年 6 月 30 日收盘时的策略净值为 1.4194。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

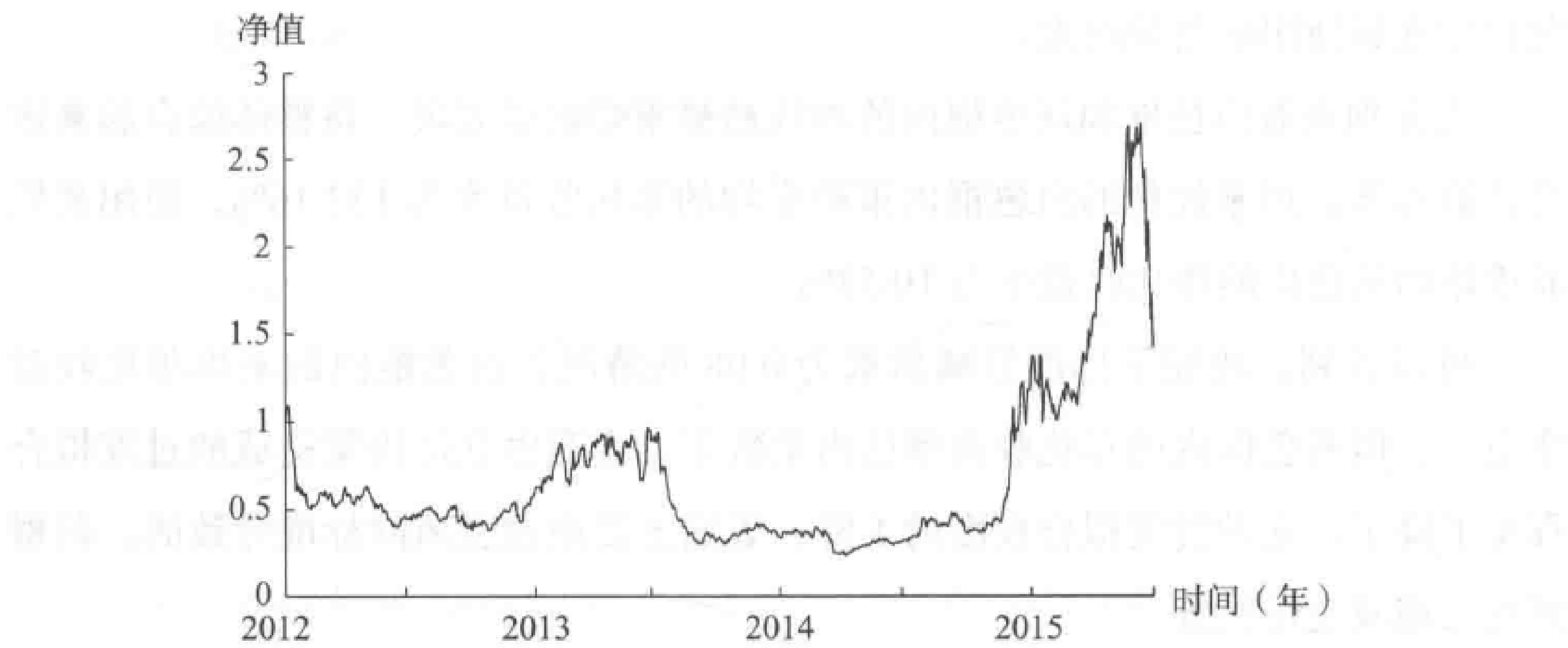


图 11-10 推进分析下仓位优化的均线趋势策略净值走势

基于整个交易过程来看，三年半的策略收益率为 41.94%，年化收益率为 10.53%，说明该仓位设置下的均线趋势策略存在一定的盈利能力。这是与不进行仓位优化或者将标准差阈值设为 0.08 时的均线趋势策略的最大区别所在。

前文说过，当标准差阈值设置为 0.08 时，使用既定仓位决策方法得到的仓位设置过重。这里将标准差阈值设置为 0.04，仓位设置随之下降，盈利能力就增加了，是一个实际的例证。

在整个三年半的模拟交易中，策略净值的最大回撤率为 78.56%，回撤的发生从 2012 年 1 月 5 日收盘的 1.0901 开始，一直到 2014 年 4 月 9 日收盘时的 0.2338 结束，整个回撤过程一共经历 543 个交易日，与第 7 章 7.1 节中的策略最大回撤期完全一致，时间较长。

经过计算，均线趋势策略在整个交易过程中平均的年化最大回撤率为 50.53%。对比第 7 章中不进行仓位决策时策略的年化最大回撤率 20.25%，风险有所升高。这是仓位加重所带来的几乎必然的结果。

11.1.3 标准差阈值 0.02

当标准差的阈值取为 0.02 时，我们基于推进分析中每一行的白色框数据，分别计算出使得收益率序列标准差等于 0.02 的仓位比例设置 $f_{风险}$。图 11-11 展示了这一长度为 845 的仓位比例时间序列。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

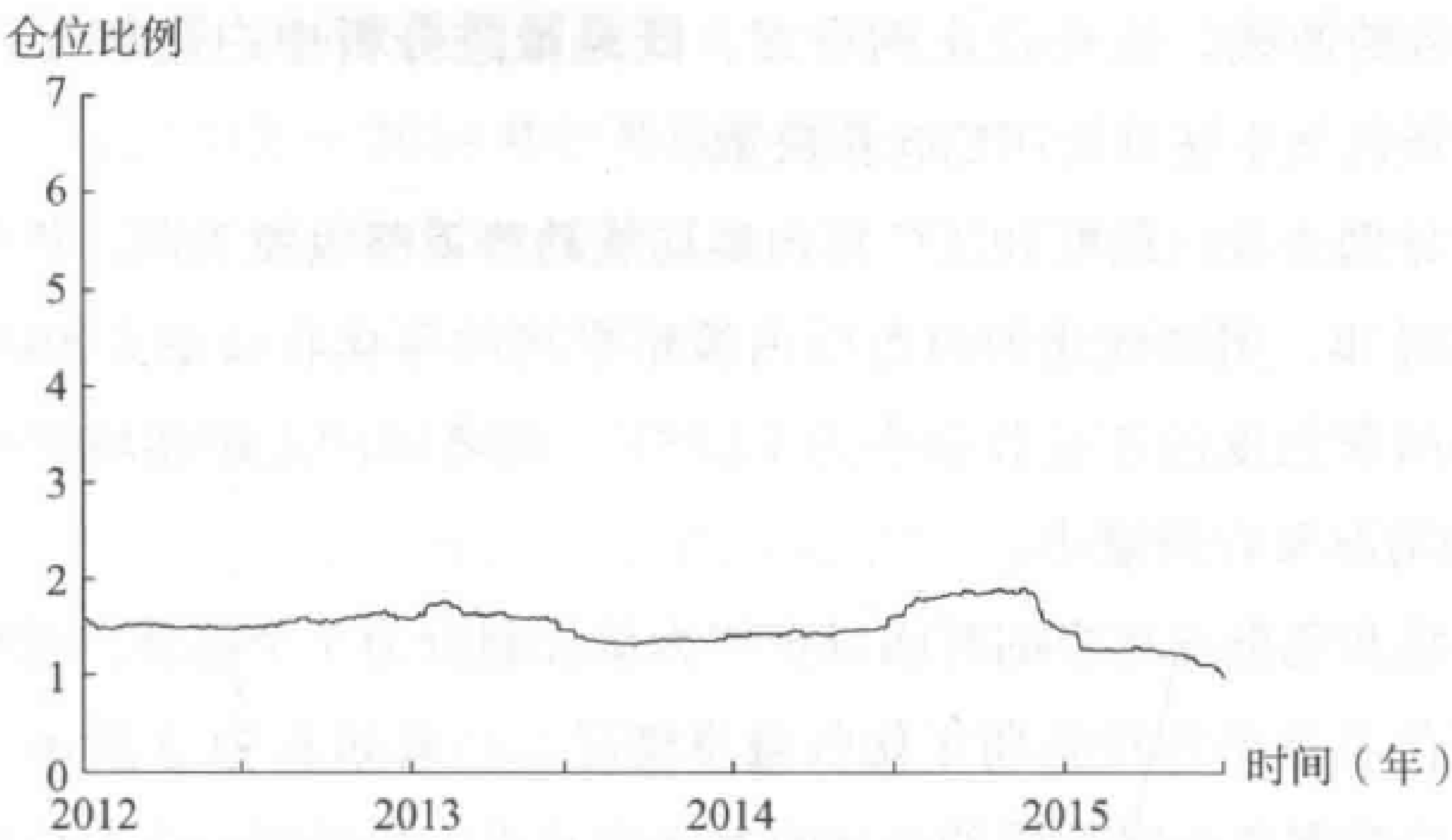


图 11-11 风险度量指标控制下的仓位设置

从图 11-11 中可以看到，当使用标准差作为风险度量指标，并限定其阈值水平为 0.02 时，仓位比例设置的运行相较于前面两种情况更为平缓。最高值为 1.8962，最低值为 0.9784，整体的仓位比例平均值为 1.4999。

接着，我们把使用凯利公式与风险度量指标得到的仓位比例时间序列放在一张图中加以对比，实际上就是将图 11-3 和图 11-11 重叠，如图 11-12 所示。可以看到，在所有时间点上，通过限定风险度量指标数值得到的仓位设置都要小于凯利公式计算得到的仓位设置，而且两者相差较大。

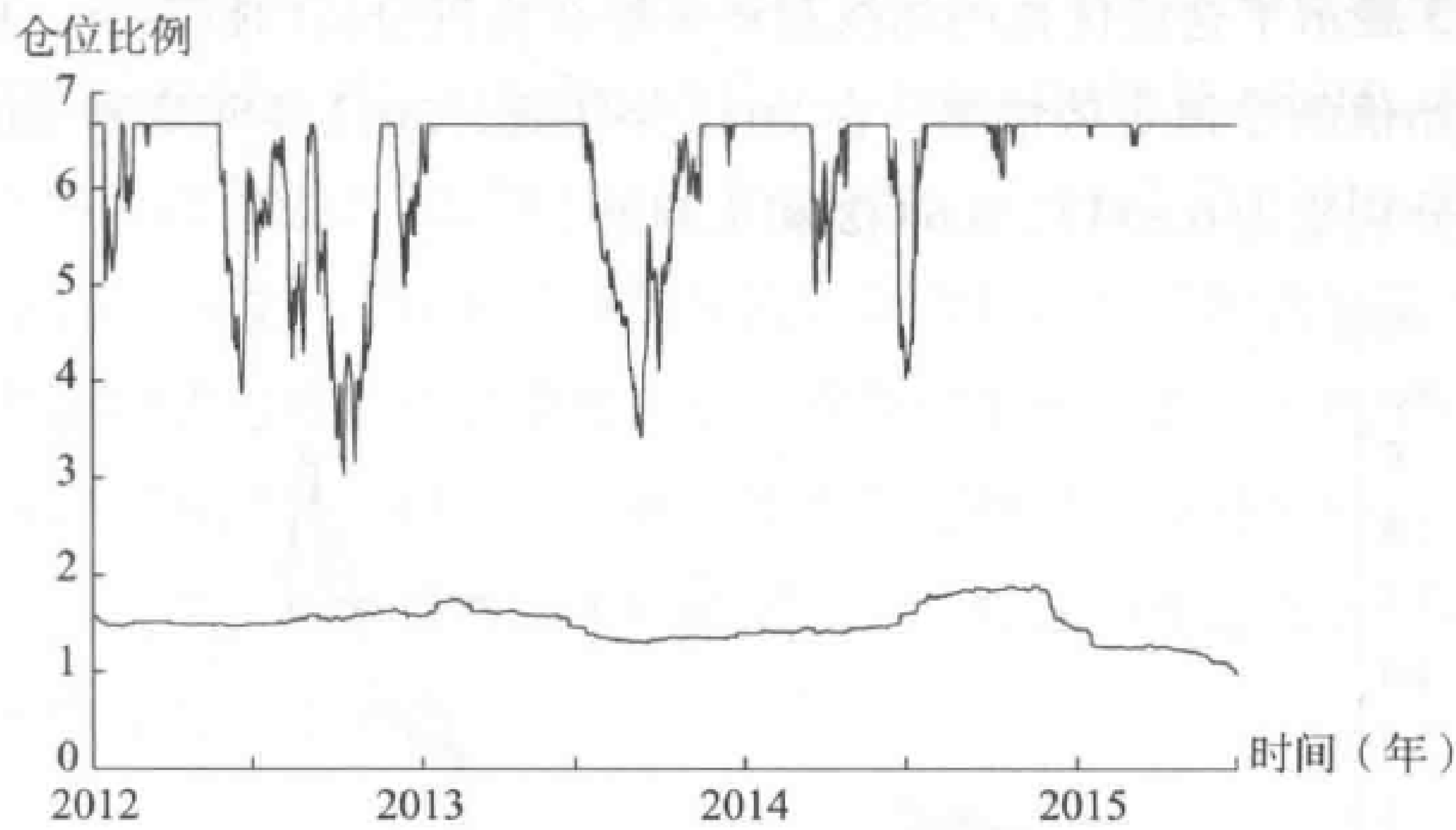


图 11-12 凯利公式与风险度量指标分别得到的仓位设置

那么当我们使用求取两种仓位决策较小值的方法得到最终结果时，最后使用的仓位设置实际上就是基于风险度量指标限制所得到的仓位设置，也就是

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

图 11-11 所示的曲线。该仓位比例设置，既是推进分析中白色框逐行优化的结果，也是在灰色框中逐日使用的仓位决策。

首先，分别查看白色框和灰色框内的均线趋势策略收益表现。将整体综合起来进行计算得知，用来优化的白色框内策略平均的年化收益率为 59.91%，而用来检验策略的灰色框的年化收益率为 12.18%。两者间的差距相较于标准差阈值取为 0.04 的情况又有所缩小。

其次，将灰色框三年半的时间以半年为单位划分为 7 个部分，考察每半年中的优化期年化收益率和检验期年化收益率情况，结果如表 11-3 所示。与表 11-2 对比来看，检验期各个时间段的年化收益率变动进一步减缓，由仓位决策导致的过度拟合程度也相应地减小了。

表 11-3 优化期年化收益率与检验期年化收益率对比 (%)

时间段	优化期年化收益率	检验期年化收益率
2012 年上半年	49.42	−53.00
2012 年下半年	35.88	47.88
2013 年上半年	69.58	50.86
2013 年下半年	47.44	−60.36
2014 年上半年	55.51	−2.85
2014 年下半年	55.61	345.96
2015 年上半年	117.64	24.16

图 11-13 展示了仓位优化的均线趋势策略在推进分析的框架下，模拟交易过程以 1 为初始值的净值走势情况。在 2012 年年底、2013 年年底和 2014 年年底，策略的净值分别变为 0.8337、0.6448 和 1.3420。

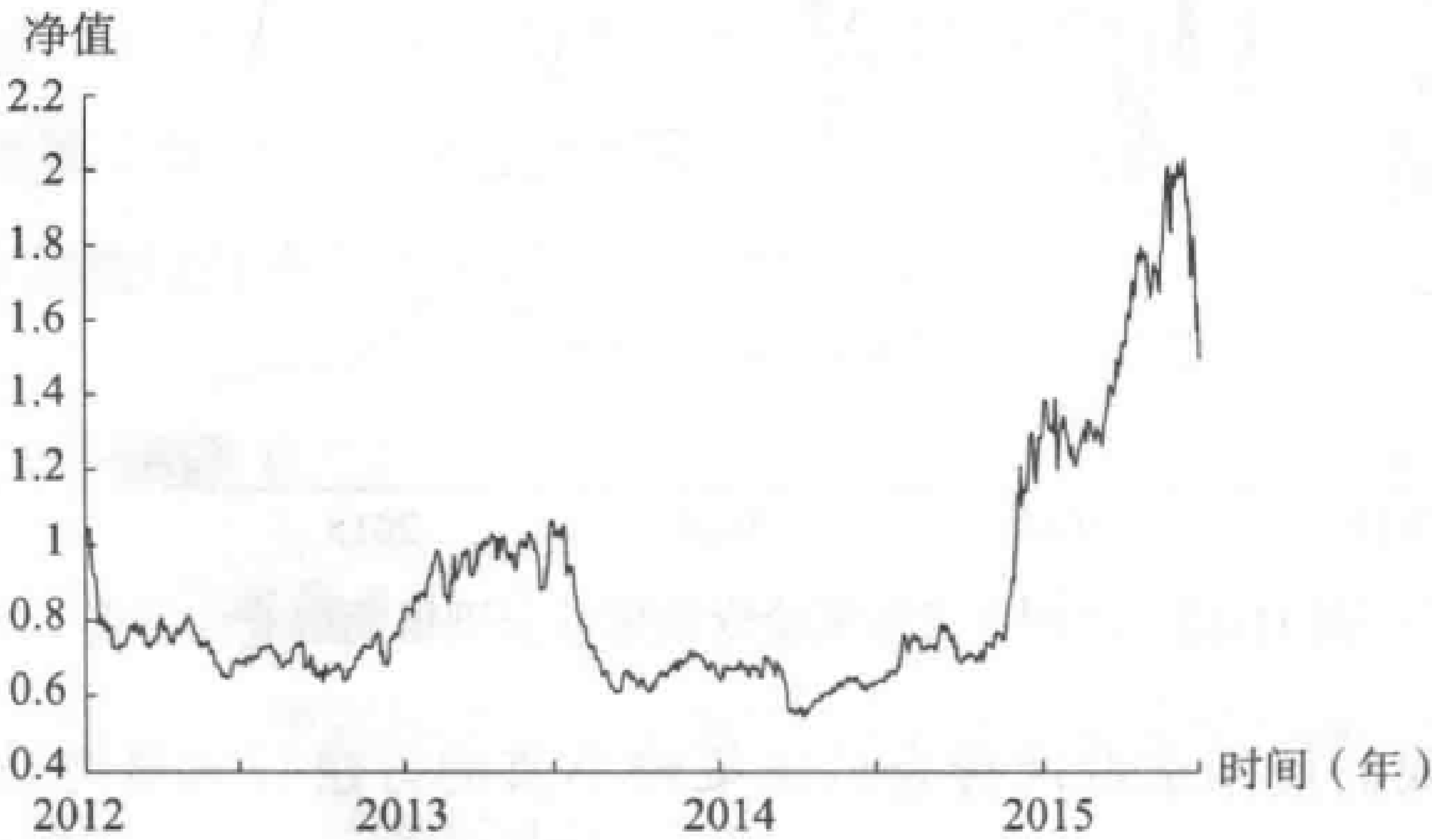


图 11-13 推进分析下仓位优化的均线趋势策略净值走势

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

净值整体走势与标准差阈值设为 0.04 时的策略净值走势基本一致，只是波动范围更小。在 2012 ~ 2014 年三年间，策略净值水平在大部分时间内都小于初始净值 1。净值从 2014 年年末开始呈现出一个上升的态势，在 2015 年 6 月 12 日达到最高点 2.0292，最后一段时间有一个较为明显的下跌，到 2015 年 6 月 30 日收盘时的策略净值为 1.4954。

基于整个交易过程来看，三年半的策略收益率为 49.54%，年化收益率为 12.18%，说明该仓位设置下的均线趋势策略存在一定的盈利能力。这一收益率水平略高于标准差阈值设为 0.04 时的策略收益。

在整个三年半的模拟交易中，策略净值的最大回撤率为 48.96%，回撤的发生从 2013 年 6 月 25 日收盘的 1.0650 开始，一直到 2014 年 4 月 9 日收盘时的 0.5436 结束，整个回撤过程一共经历 192 个交易日。

经过计算，均线趋势策略在整个交易过程中平均的年化最大回撤率为 29.71%，较之标准差阈值 0.04 时的 50.53% 有了较大幅度的下降。

最后，我们尝试性地使用凯利公式来计算一下，如果在灰色框检验期内使用同一个比例的仓位设置，那么这个固定比例如何设置才能使得最终的策略盈利能力最大。这一结果在实践中是没有意义的，因为这样做一定存在未来信息的问题。但是对比该结果，有助于我们对之前所做的仓位决策进行思考。

具体做法就是基于第 7 章 7.1 节中没有仓位决策下的策略检验期收益率序列，通过凯利公式计算得到对应的最优仓位比例，最终的结果是 1.3517。而当标准差阈值取为 0.02 时，动态仓位设置的平均值为 1.4999，是比较接近于这一固定数值的。

也就是说，通过一个比较严格的风险控制级别，得到的仓位结果反而更接近于能使得其盈利能力最大化的设置。这个现象产生的根源，就在于优化期与检验期之间的过度拟合问题。研究人员有必要在实际的量化交易策略研发过程中，仔细区别这两类情况，不要只着眼于回溯测试中的较好的收益结果，进而丧失掉了对真正的风险控制的把握。

11.2 仓位优化的自回归策略

上一节中，作者在第 7 章推进的均线趋势策略的基础上，加入了第 10 章所

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

述的仓位决策模型方法，将买卖和仓位两者相结合，构造了一个更为完善的量化交易策略。我们曾在第 4 章 4.1 节里将择时策略框架大致分为分类和预期两种。上一节中的均线趋势策略基于分类形成，这一节，我们就将针对预测类的自回归策略，结合第 7 章和第 10 章的内容，再进行一次融合仓位决策的策略构造。

首先论述的是第 7 章 7.4 节所述的自回归策略。推进分析的框架依然保留。自回归策略的具体设置仍然与第 4 章 4.4 节中保持一致，通过前 9 日的日度收益率回归得到交易时刻的日度收益率，基于收益率预测值的情况给出多空的看法，即交易时刻的日度收益率预测值为正时则看多，为负时则看空。具体可参见图 4-9 和相关叙述。

数据方面也和前面所选取的数据一致，收集得到沪深 300 股指期货主力合约从 2010 年 4 月 19 日到 2015 年 6 月 30 日的日收盘价数据。与上一节的设置一样，基于样本时间段内的期货市场情况，将股指期货的交易保证金设为 15%。

大的推进分析框架仍然维持第 7 章 7.4 节中的设定。与上一节类似，在每一行的白色框优化期中，不仅要进行自回归模型的参数估计，进而根据这一自回归模型的预测值正负号来判断涨跌，还要基于所得到的收益率序列，计算出相应的最优仓位设置，最后将自回归模型和仓位设置同时应用于对应的灰色框上。

以第一行操作为例，白色框为优化期，从 2010 年 12 月 31 日的收盘开始，直到 2011 年 12 月 30 日的收盘结束，一共 244 个交易日。这 244 个交易日的日度收益率与各自之前的 9 个日度收益率组成了 244 组数据，针对这 244 组数据进行自回归模型的估计，得到相应的参数估计结果。

回归参数固定之后，再分别使用之前的 9 个日度收益率，基于回归模型计算出这 244 个交易日的预期收益率。如果某一日的预期收益率为正，则该日看多，策略收益率就是这一日的日度收益率；如果某一日的预期收益率为负，则该日看空，策略收益率就是这一日的日度收益率乘以 −1。这样，在白色框的时间段上，可以获得 244 个交易日的策略收益率。

在得到白色框中不考虑仓位优化的、长度为 244 的策略收益率序列之后，我们将每一日的交易看作一次交易，套用第 10 章中同时考虑凯利公式与风险度量指标的方法，就可以得出策略在优化期内的最优仓位设置了，如下：

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

$$f=\min\{f_{凯利}，f_{风险}\}$$

在基于 2010 年 12 月 31 日到 2011 年 12 月 30 日的数据得到参数拟合的回归模型以及仓位设置之后，就可以由白色框转入灰色框进行检验。

使用估计好的自回归模型，以及 2011 年 12 月 19 日收盘到 2011 年 12 月 30 日收盘之间的 9 个日度收益率数据，来预测 2011 年 12 月 30 日收盘到 2012 年 1 月 4 日收盘的收益率。该收益预期为正时，就对 2011 年 12 月 30 日收盘到 2012 年 1 月 4 日收盘的走势看多；该收益预期为负时，就对 2011 年 12 月 30 日收盘到 2012 年 1 月 4 日收盘的走势看空。

2011 年 12 月 30 日收盘到 2012 年 1 月 4 日收盘期间的仓位，则直接使用前面优化得到的仓位比例设置，看涨时仓位为多，看跌时仓位为空。也就是说，仓位的方向和仓位的比例，共同构成了整体的仓位决策结果。

在灰色框的建仓时点，也就是 2011 年 12 月 30 日收盘时，根据上一个交易日的仓位情况，计算其与目前的仓位决策的差值，就是当前所需要进行的具体交易数额。这一过程与上一节中所描述的操作一致。

以上的这一系列操作就是图 11-1 中第一行所表现的具体流程。完成之后，将白色框和灰色框都分别向后推进一日，重复这种样本外检验的过程，直到最后一行为止。每一个灰色框也就是每一个交易日之间，都存在着重新估计回归参数、计算仓位决策、进行买卖操作从而满足新的仓位要求的过程。

最后一行中被检验的灰色框为 2015 年 6 月 29 日收盘到 2015 年 6 月 30 日收盘，也就是触及样本数据的最后一个数据。总的来看，在推进分析的过程中，模拟买卖的行为从 2011 年 12 月 30 日收盘时开始，到 2015 年 6 月 30 日收盘时结束，一共交易三年半的时间，也就是所有灰色框所覆盖的区间。

上述操作流程，我们通过图 11-14 进行展示。整体流程与上一节中均线趋势策略的推进分析差异不大，主要的差别在于判断多空时使用的模型策略不同，一个是回归模型，一个是双均线趋势判断。

与此同时，自回归策略还存在另外一种仓位比例的计算方法，即完全使用回归模型预测出的收益率序列来计算最优仓位设置，如图 11-15 所示。但是一般而言我们并不推荐这种处理方法，原因在于，预测出的收益率往往与现实情况有较

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

大的偏误，通过其优化得出的仓位结果的可信程度就比较低了。

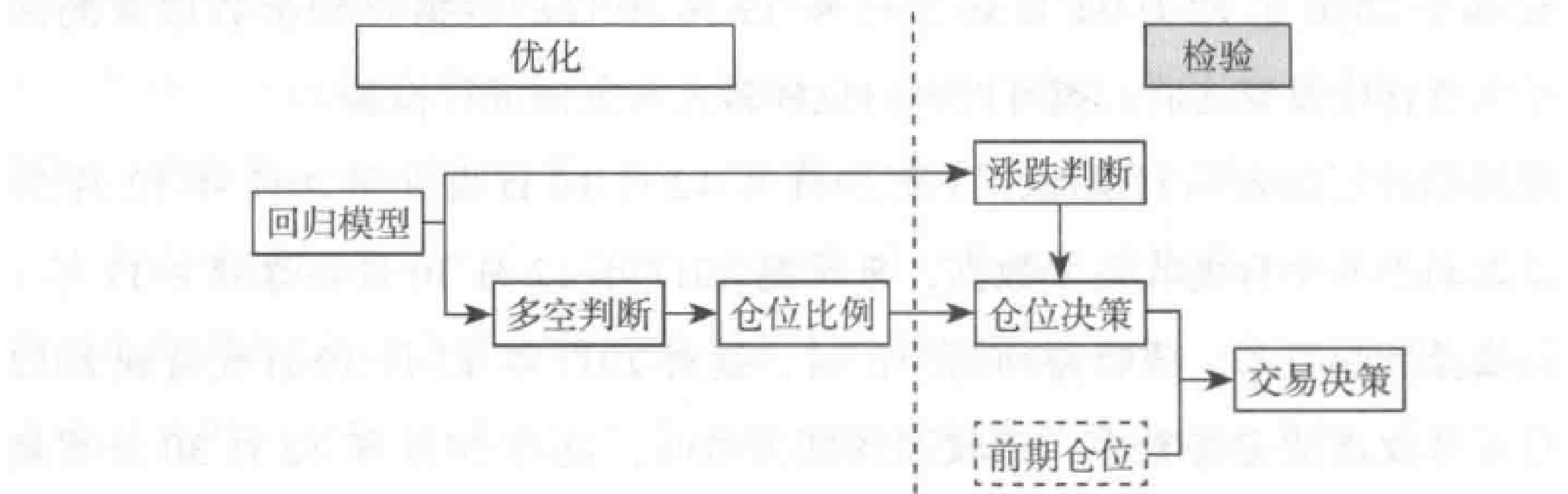


图 11-14　推进分析每一行中优化和检验过程的示意图

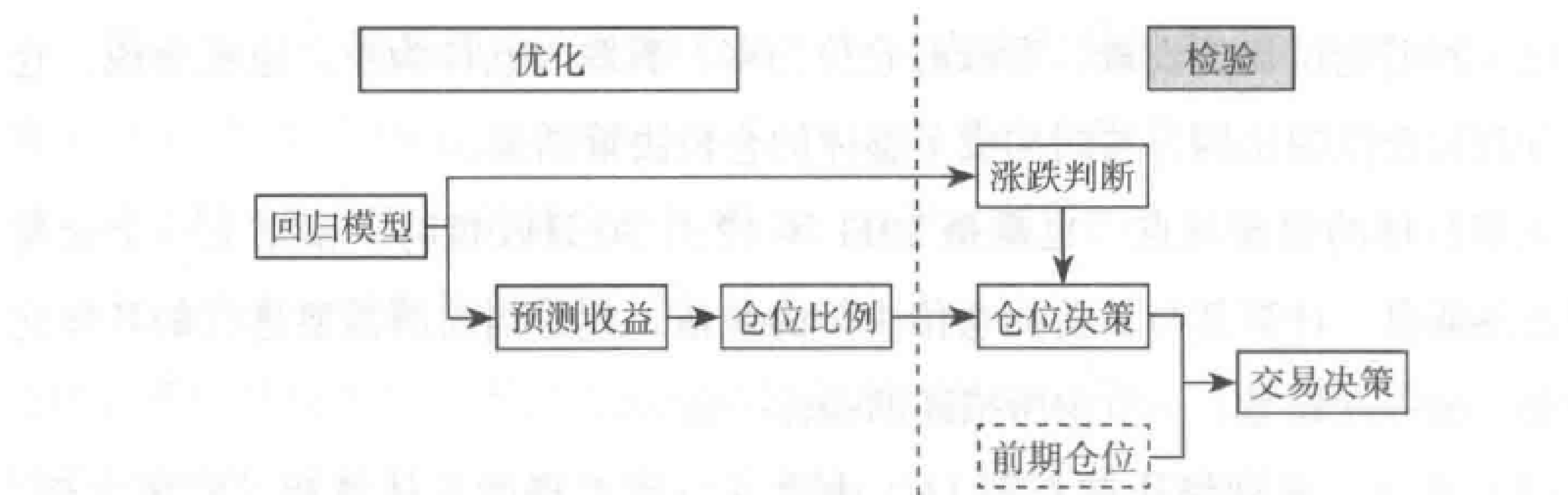


图 11-15　另一种推进分析优化和检验过程的示意图

回溯测试方面，仍然首先对凯利公式所导出的仓位结果做统一观察。图 11-16 描述了基于推进分析中每一行的白色框数据，分别使用凯利公式计算得到的最优的仓位比例设置，即 $f_{凯利}$，对应到灰色框中就是相应的仓位决策结果。推进分析一共进行了 845 行，图 11-16 所展示的也就是一个长度为 845 的时间序列。由于股指期货的交易保证金设为 15%，因此仓位比例的可选择范围就是 0 ～ 6.6667。

相比均线趋势策略下的图 11-3，这里的仓位设置在更多的时间内达到其上限 6.6667。仓位比例序列只在 2015 年上半年中存在低于上限的情况，最低点是 2015 年年初的 5.4810。这种变化表现与图 11-3 是不一样的，最低点的数值也要高于图 11-3 中的 3.0379。

一般而言，盈利越平稳，凯利公式计算得到的最优仓位比例也就越高。因此，自回归策略在凯利公式下相对更高的仓位设置结果，主要来源于其在白色框

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

优化期中更好的收益率表现。更进一步说，由于白色框中的收益率受到模型拟合的影响，因此较高的仓位结果来源于自回归模型这一复杂程度更高的模型设置。

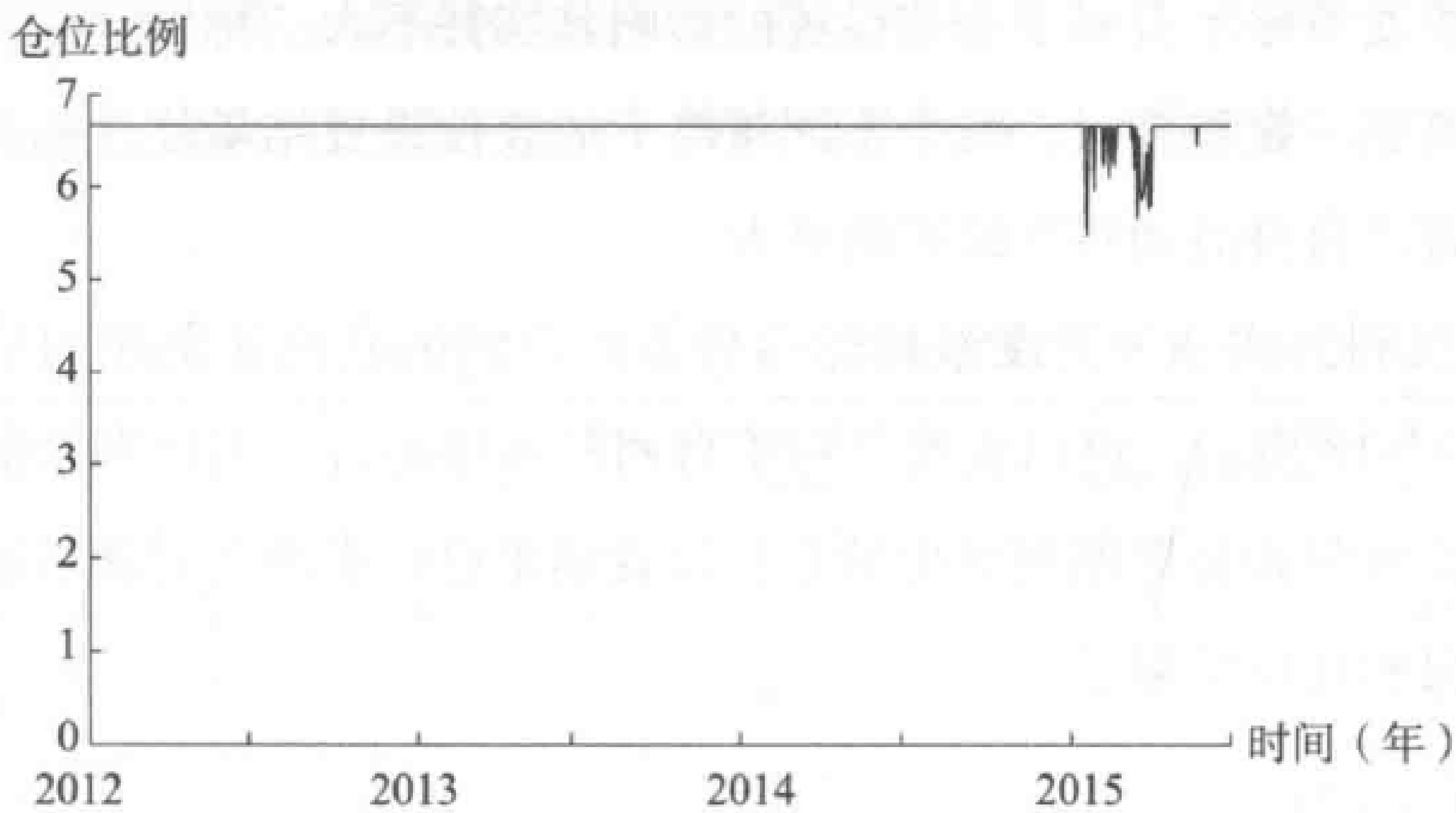


图 11-16 凯利公式下的仓位设置

然后，我们进一步加入风险度量指标的限定。与上一节中的做法一样，将风险度量指标设定为标准差，阈值水平则分别设为三个不同的数值，即 0.08、0.04 和 0.02，分别进行叙述。

11.2.1 标准差阈值 0.08

当标准差的阈值取为 0.08 时，基于推进分析中每一行的白色框数据，分别计算出使得收益率序列标准差等于 0.08 的仓位比例设置，即 $f_{风险}$。这也是一个长度为 845 的时间序列，如图 11-17 所示。

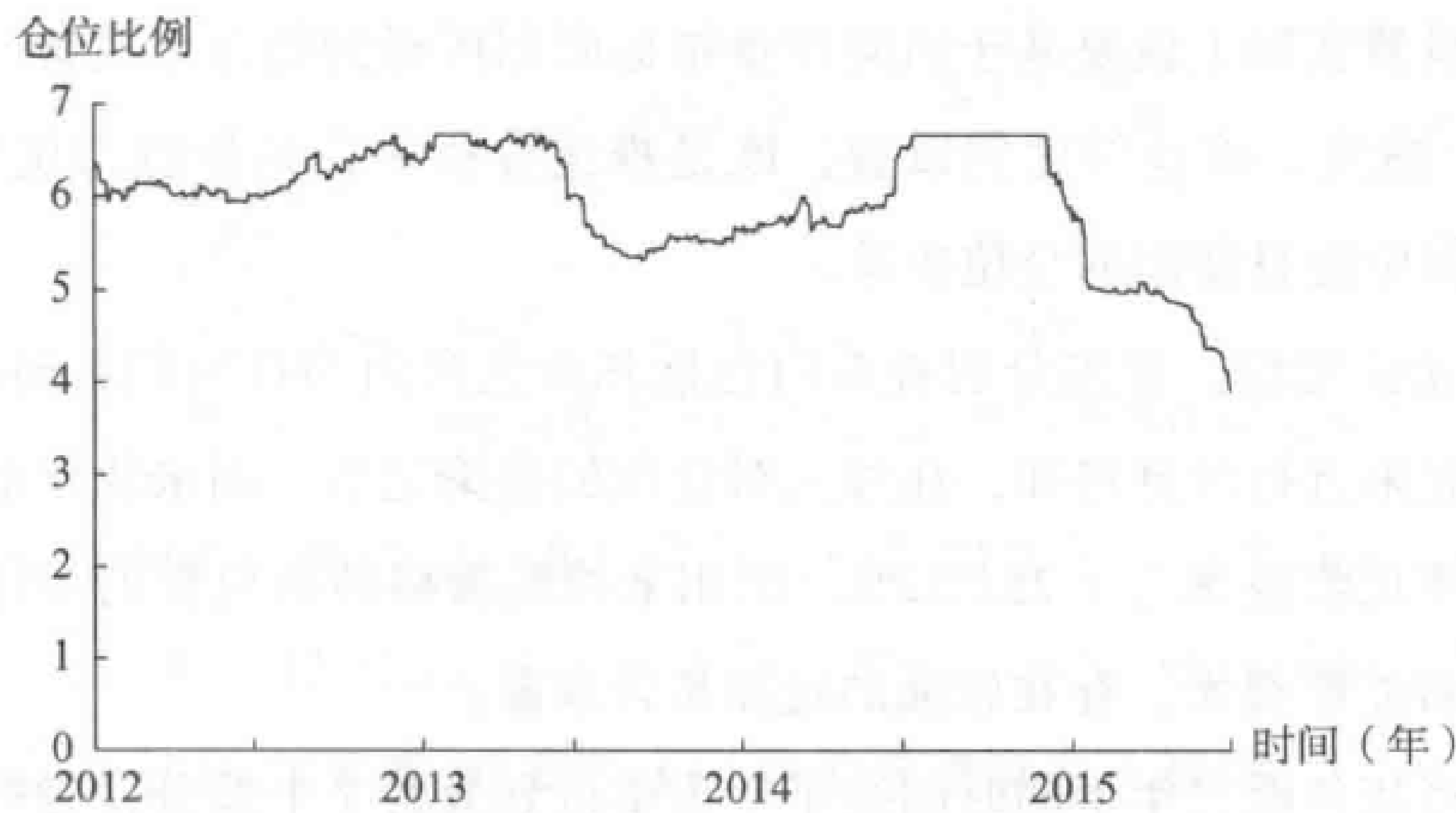


图 11-17 风险度量指标控制下的仓位设置

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

可以先将其与图 11-4 进行对比。实际上，两幅图除了一些细节上的细微差别以外，相似程度非常高。这说明当我们使用控制风险度量指标的方式来进行仓位决策时，买卖策略本身对于仓位设置的影响其实并不大。特别是涨跌判断的准确率并没有高到一定程度时，两个不同策略下的仓位设置结果往往相近，仓位决策受到交易资产自身波动特性的影响更大。

把使用凯利公式和风险度量指标两种方法得到的仓位比例序列放在同一张图中，如图 11-18 所示。可以发现，对于自回归策略而言，当标准差的阈值设为 0.08 时，风险度量指标限制所导出的仓位设置结果已经在所有时间点都小于凯利公式计算得出的仓位结果了。

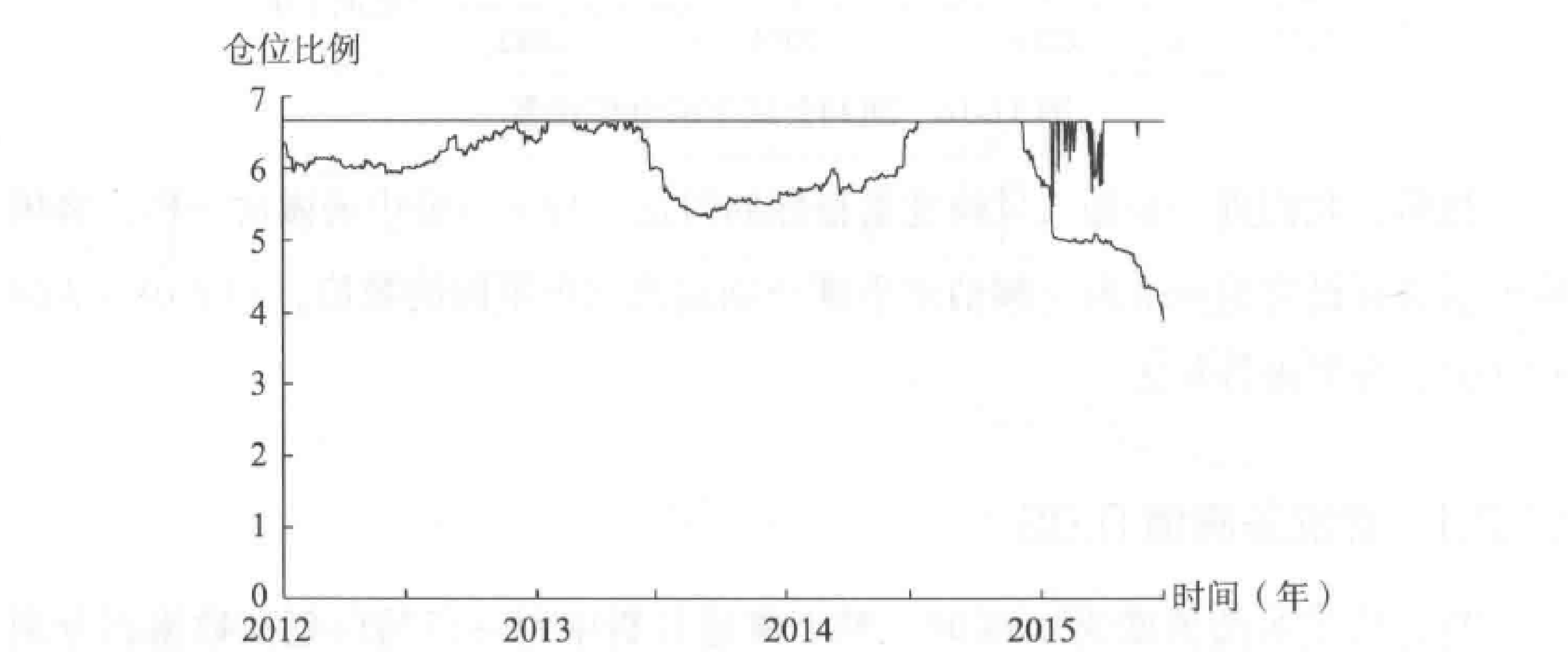


图 11-18 凯利公式与风险度量指标分别得到的仓位设置

那么，当我们使用求取两种仓位决策较小值的方法得到最终结果时，最后使用的仓位设置实际上就是基于风险度量指标限制所得到的仓位设置，也就是图 11-17 所示的曲线。该仓位比例设置，既是推进分析中白色框逐行优化的结果，也是在灰色框中逐日使用的仓位决策。

设置好仓位之后，首先分别查看白色框和灰色框内的自回归策略收益表现。将整体综合起来进行计算得知，在加入对仓位的决策之后，用来优化的白色框内策略平均的年化收益率为 1 252.52%，而用来检验策略的灰色框的年化收益率为 50.84%。两者差距很大，存在较强的过度拟合现象。

其次，将灰色框三年半的时间按半年为单位划分为 7 个部分，考察每半年中

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

的优化期年化收益率和检验期年化收益率情况，用以说明具体时间内参数的过度拟合情况，结果如表 11-4 所示。值得注意的是，两个收益率之间差距的变化与表 11-1 中的情况存在一定的相似性，且都在 2014 年下半年表现出检验期收益高于优化期收益的情况。这或许暗示了均线趋势策略与自回归策略之间某种内在的联系，值得进一步思考。

表 11-4 优化期年化收益率与检验期年化收益率对比 (%)

时间段	优化期年化收益率	检验期年化收益率
2012 年上半年	1 113.26	21.04
2012 年下半年	892.19	996.50
2013 年上半年	2 293.81	87.63
2013 年下半年	3 397.31	44.32
2014 年上半年	1 098.73	−59.40
2014 年下半年	995.73	1 498.35
2015 年上半年	525.46	−92.38

图 11-19 展示了仓位优化的自回归策略在推进分析的框架下，模拟交易过程以 1 为初始值的净值走势情况。在 2012 年年底、2013 年年底和 2014 年年底，策略的净值分别变为 3.6431、5.9949 和 15.2715。

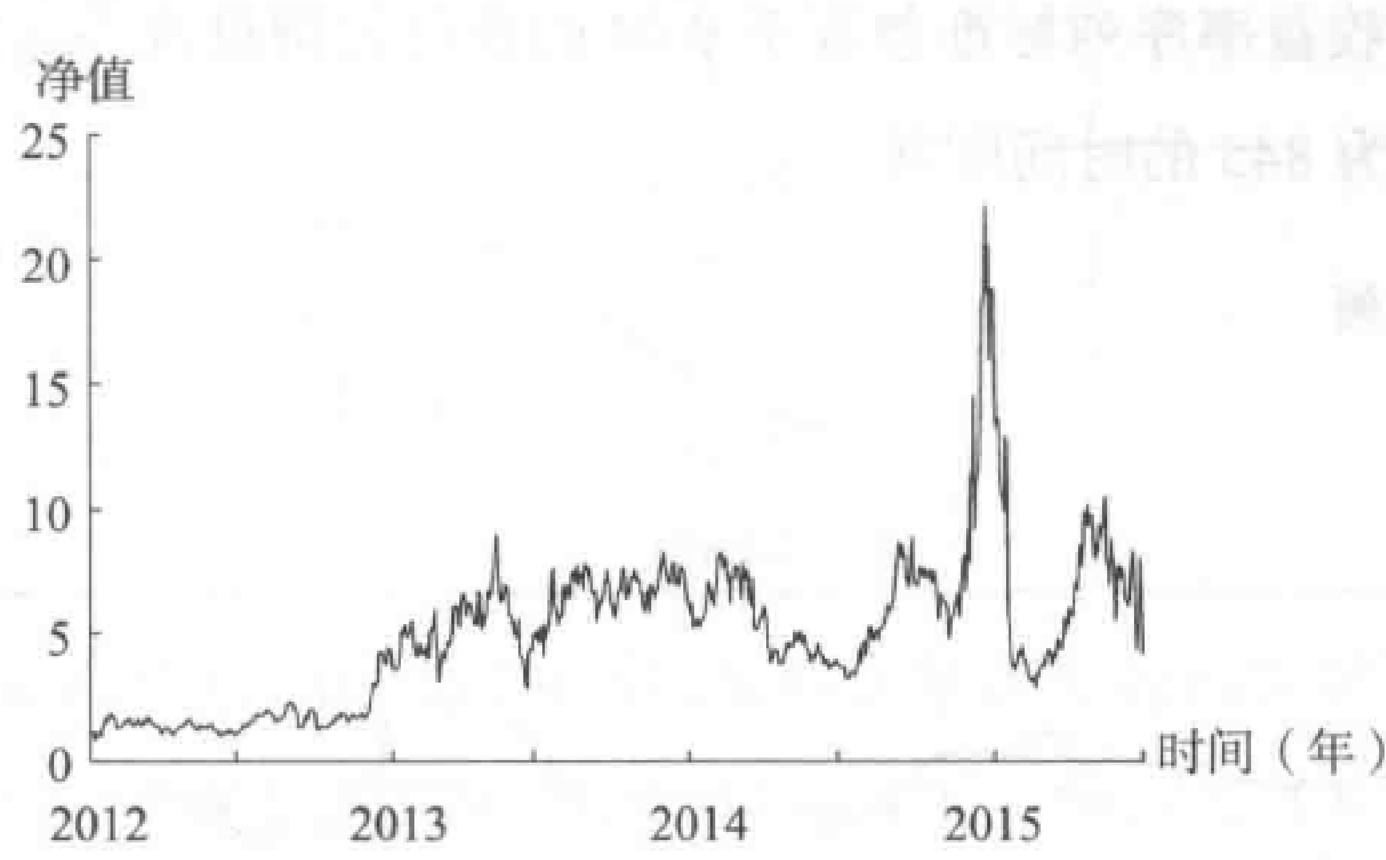


图 11-19 推进分析下仓位优化的自回归策略净值走势

在三年半的检验期当中，仓位优化后的策略净值基本都在初始净值 1 之上。净值水平于 2014 年 12 月 22 日达到最高，为 22.1566。随后净值又有一个极大的下降，最后到 2015 年 6 月 30 日收盘时的策略净值为 4.2151。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

基于整个交易过程来看，三年半的策略收益率为321.51%，年化收益率为50.84%，说明该仓位设置下的自回归策略具备较强的盈利能力。参考第7章7.4节的结果，当不进行仓位决策，始终保持100%的仓位比例时，策略同期的年化收益率为21.53%。两相对比，仓位设置为收益的进一步提升提供了一定的帮助。

在整个三年半的模拟交易中，策略净值的最大回撤率为86.83%。回撤的发生从2014年12月22日收盘的22.1566开始，一直到2015年2月25日收盘时的2.9179结束。整个回撤过程一共经历40个交易日，与第7章7.4节中的策略最大回撤期完全一致，是一个相对而言较为急促的回撤过程。

经过计算，自回归策略在整个交易过程中平均的年化最大回撤率为65.40%。对比来看，第7章7.4节中不进行仓位决策的自回归策略年化最大回撤率仅为14.52%。仓位决策的引入、仓位比例的加重，极大地提升了策略的风险。

11.2.2 标准差阈值0.04

当标准差的阈值取为0.04时，我们基于推进分析中每一行的白色框数据，分别计算出使得收益率序列标准差等于0.04的仓位比例设置$f_{风险}$，并在图11-20中展示这一长度为845的时间序列。

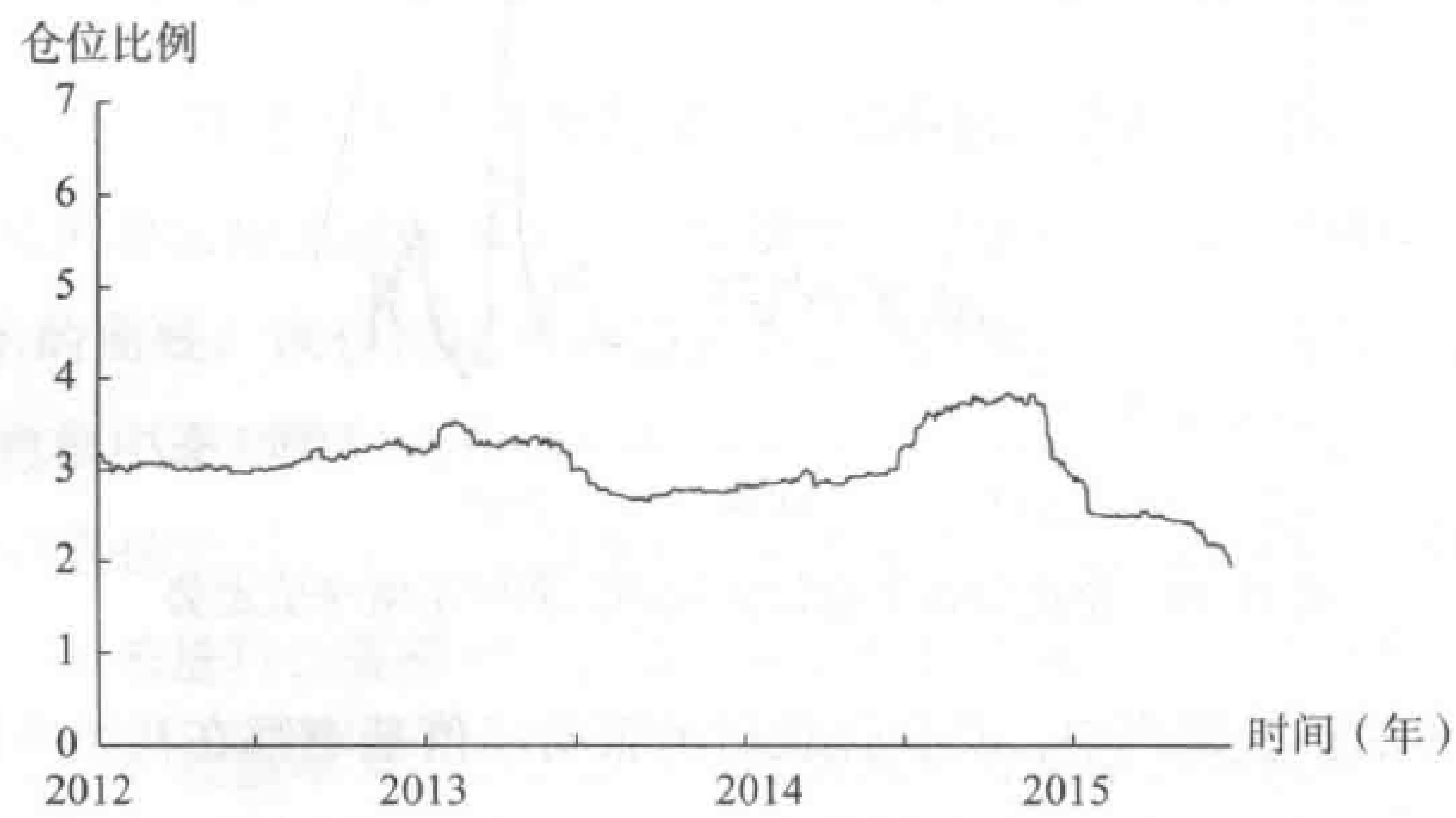


图 11-20 风险度量指标控制下的仓位设置

很明显，在所有时间点上，通过限定风险度量指标数值得到的仓位设置都要

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

小于凯利公式计算得到的仓位设置。因此，当我们使用求取两种仓位决策较小值的方法得到最终结果时，最后使用的仓位设置实际上就是图 11-20 所示的仓位比例。该仓位比例设置，既是推进分析中白色框逐行优化的结果，也是在灰色框中逐日使用的仓位决策。

设置好仓位之后，首先分别查看白色框和灰色框内的自回归策略收益表现。将整体综合起来进行计算得知，用来优化的白色框内策略平均年化收益率为 348.34%，而用来检验策略的灰色框的年化收益率为 61.96%。

这一结果相较于标准差阈值取为 0.08 时的情况，白色框内的平均年化收益率下降了很多，但灰色框内的年化收益率水平反而上升。有理由判断，随着仓位比例的下降，由仓位决策所导致的过度拟合现象有了一定的缓解。

其次，将灰色框三年半的时间按半年为单位划分为 7 个部分，考察每半年中的优化期年化收益率和检验期年化收益率情况，结果如表 11-5 所示。与表 11-4 中相似，检验期收益率比优化期收益率更高的仍是 2012 下半年和 2014 下半年。

表 11-5 优化期年化收益率与检验期年化收益率对比 (%)

时间段	优化期年化收益率	检验期年化收益率
2012 年上半年	319.85	32.68
2012 年下半年	285.75	298.79
2013 年上半年	489.45	80.56
2013 年下半年	617.00	42.22
2014 年上半年	314.64	−26.80
2014 年下半年	329.62	518.44
2015 年上半年	198.60	−52.48

图 11-21 展示了仓位优化的自回归策略在推进分析的框架下，模拟交易过程以 1 为初始值的净值走势情况。在 2012 年年底、2013 年年底和 2014 年年底，策略的净值分别变为 2.3003、3.6861 和 7.8427。

相比图 11-19，图 11-21 中的净值走势在趋势上十分类似，但是更为平滑。净值水平同样于 2014 年 12 月 22 日达到最高，为 9.2461。因此，随后的净值回撤程度也相对较小。随着 2015 年 2 月之后的一段回升，到 2015 年 6 月 30 日收盘时的策略净值为 5.4066。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

基于整个交易过程来看，三年半的策略收益率为440.66%，年化收益率为61.96%。该仓位设置下的自回归策略同样存在一定的盈利能力。值得注意的是，其比标准差阈值设为0.08时的策略盈利能力更强，但是整个净值走势更为平缓，这是一个吸引研究者的特征。

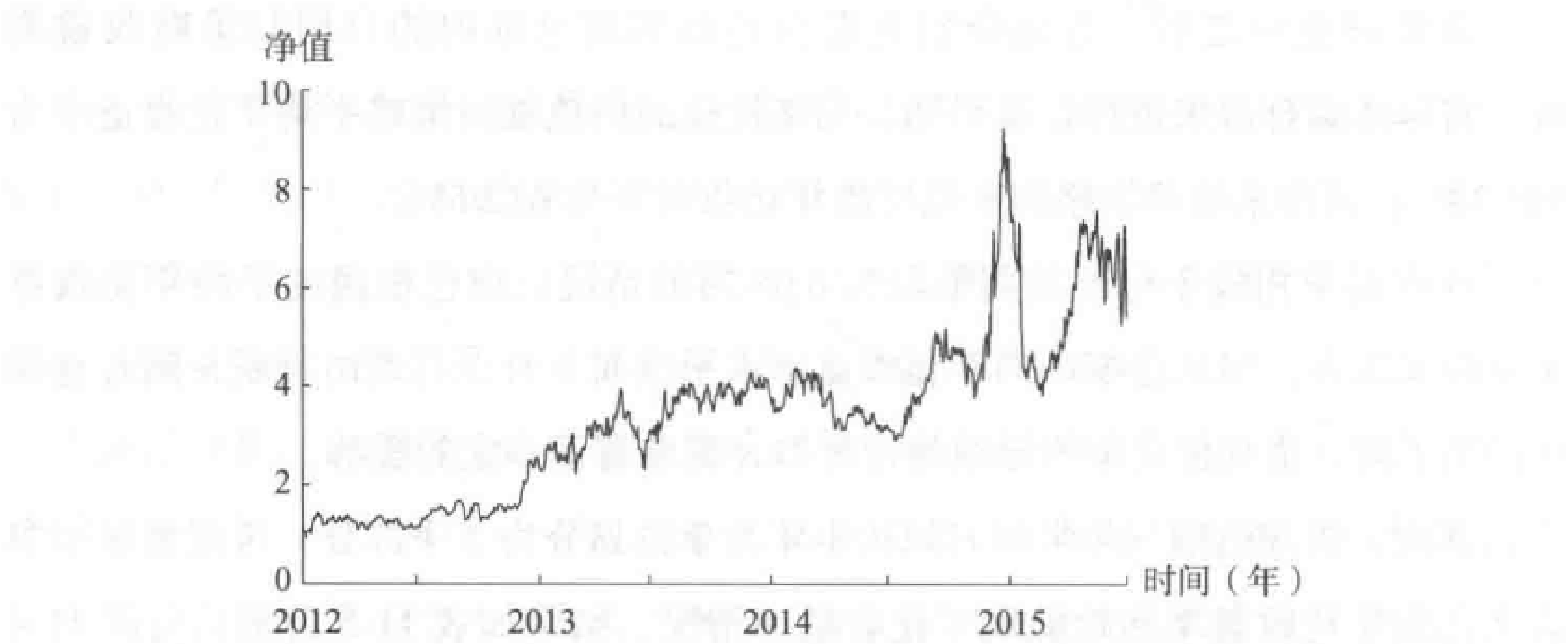


图 11-21　推进分析下仓位优化的自回归策略净值走势

在整个三年半的模拟交易中，策略净值的最大回撤率为58.74%。回撤的发生从2014年12月22日收盘的9.2461开始，一直到2015年2月25日收盘时的3.8153结束。这与之前的最大回撤期也是一致的。

经过计算，自回归策略在整个交易过程中平均的年化最大回撤率为39.29%。对比标准差阈值设为0.08时的策略年化最大回撤率65.40%，风险有所减小。与此同时，策略的收益水平反而有所增加。如果使用“收益 / 回撤”一类的度量指标来判断，将标准差的阈值设为0.04明显是一个更优的选择。

11.2.3　标准差阈值0.02

当标准差的阈值取为0.02时，我们基于推进分析中每一行的白色框数据，分别计算出使得收益率序列标准差等于0.02的仓位比例设置$f_{风险}$。图11-22展示了这一长度为845的仓位比例时间序列。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

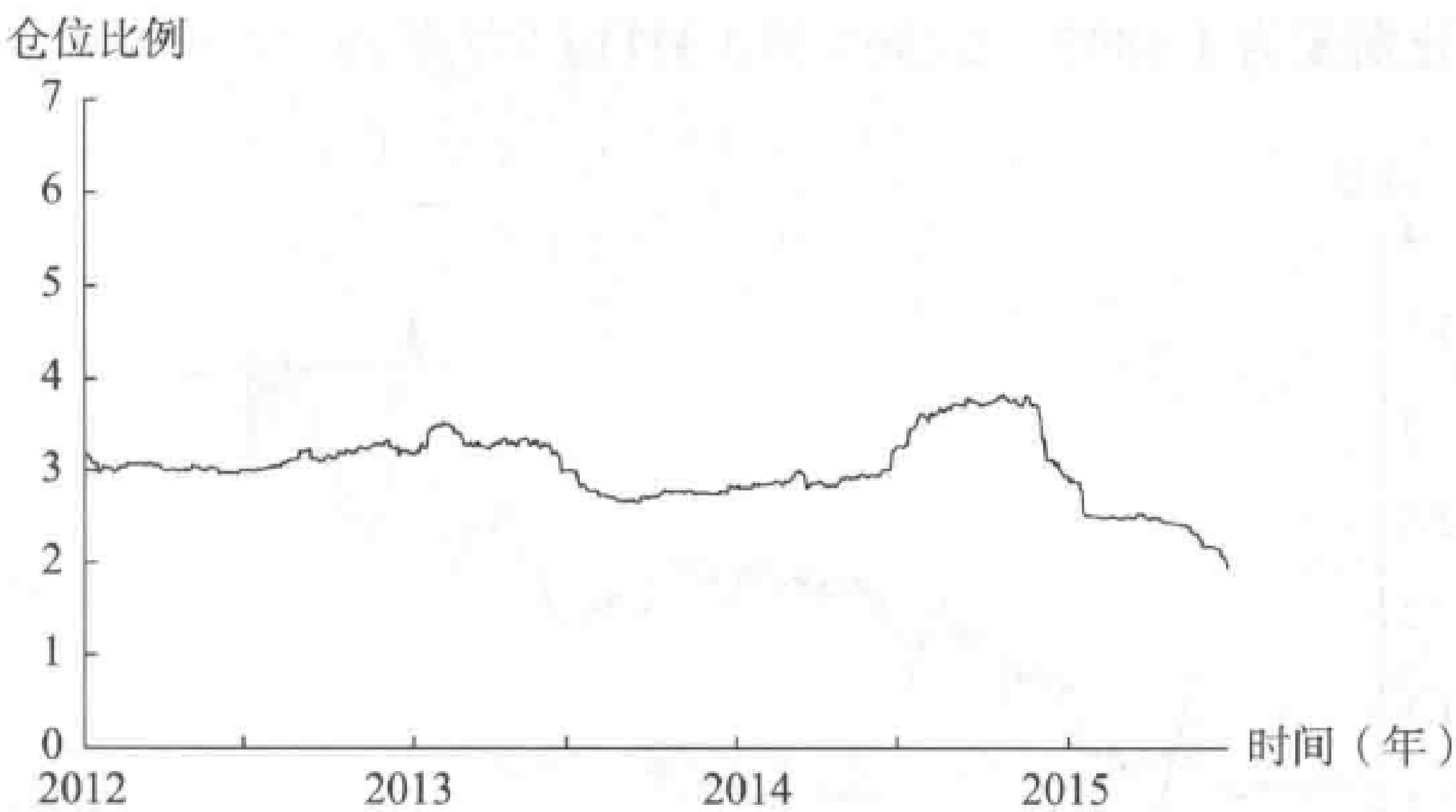


图 11-22 风险度量指标控制下的仓位设置

显然，通过限定风险度量指标得到的仓位设置始终小于凯利公式计算得到的仓位设置。因此，当我们使用求取两种仓位决策较小值的方法得到最终结果时，求得的仓位设置实际上就是图 11-22 所示的仓位比例。该仓位比例设置，既是推进分析中白色框逐行优化的结果，也是在灰色框中逐日使用的仓位决策。

首先，分别查看白色框和灰色框内的自回归策略收益表现。将整体综合起来进行计算得知，用来优化的白色框内策略平均的年化收益率为 122.15%，而用来检验策略的灰色框的年化收益率为 35.82%。标准差阈值由 0.04 变为 0.02 之后，过度拟合的问题进一步缓解。

细分来看，将灰色框三年半的时间以半年为单位划分为 7 个部分，考察每半年中的优化期年化收益率和检验期年化收益率情况，结果如表 11-6 所示。

表 11-6 优化期年化收益率与检验期年化收益率对比 （%）

时间段	优化期年化收益率	检验期年化收益率
2012 年上半年	114.78	20.72
2012 年下半年	106.62	109.28
2013 年上半年	154.17	44.40
2013 年下半年	181.44	24.43
2014 年上半年	113.34	−11.52
2014 年下半年	117.94	172.93
2015 年上半年	80.91	−22.25

图 11-23 展示了仓位优化的自回归策略在推进分析的框架下，模拟交易过程以 1 为初始值的净值走势情况。在 2012 年年底、2013 年年底和 2014 年年底，

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

策略的净值分别变为 1.5895、2.1307 和 3.3111。

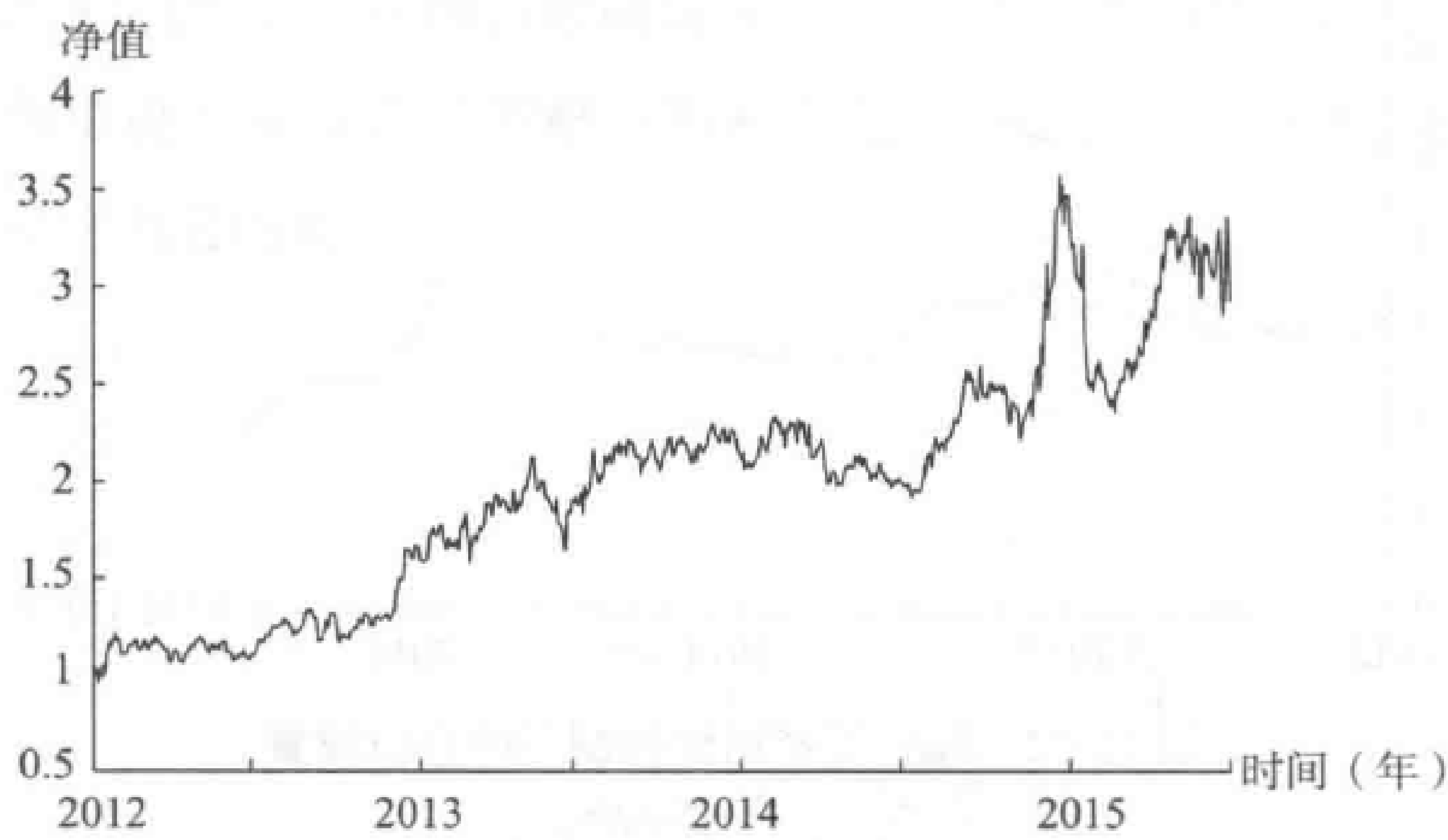


图 11-23　推进分析下仓位优化的自回归策略净值走势

净值整体走势与标准差阈值设为 0.04 时的策略净值走势基本一致，只是趋势向上的性质更加明显，走势更为平缓。净值同样在 2014 年 12 月 22 日达到最高值，数值为 3.5773，在最高点附近呈现出一个波动相对较小的倒 V 字形状。样本期末，即 2015 年 6 月 30 日收盘时的策略净值为 2.9196。

基于整个交易过程来看，三年半的策略收益率为 191.96%，年化收益率为 35.82%。说明该仓位设置下的自回归策略也是具有一定盈利能力的。只不过，这一收益率水平要低于标准差阈值设为 0.04 时的策略整体收益。

在整个三年半的模拟交易中，策略净值的最大回撤率为 34.09%。回撤的发生从 2014 年 12 月 22 日收盘的 3.5773 开始，一直到 2015 年 2 月 25 日收盘时的 2.3579 结束。最大回撤期同样与之前的一致。

经过计算，自回归策略在整个交易过程中平均的年化最大回撤率为 21.48%。对比标准差阈值设为 0.04 时的策略年化最大回撤率 39.29%，风险进一步减小。与此同时，策略的年化收益率也由 61.96% 降低至 35.82%。如果同时考虑策略收益和回撤风险，标准差阈值设为 0.02 和 0.04 各有优劣。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

第 12 章

投资组合决策

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

12.1 最优投资组合理论

在第 10 章中，我们介绍了针对量化择时策略的仓位决策方法。然后，第 11 章基于所介绍的仓位决策方法，研究了加入仓位优化的择时策略的两个具体案例。依照这一思路，我们将在这一章和下一章中论述仓位优化的选股策略。

正如第 2 章中曾提到的，买卖和仓位虽然是更为通用的说法，但是更适合于描述择时策略。针对选股策略而言，“仓位”的提法换为“配比”更为准确。相应的，我们要研究的就是比例优化的股票配置策略，或者说加入了投资组合决策的选股策略。

图 12-1 通过灰色区域说明了投资组合决策在量化选股策略研发流程中的位置。如图 12-1 所示，各只股票的具体配比也就是所要做的具体的投资组合决策，取决于股票的收益与风险。与择时策略仓位决策不同的是，投资组合决策往往是收益与风险同时作用的结果，决策模型一般要求同时输入收益和风险。而本节要介绍的最优投资组合理论，正是这一类模型中的典型代表。

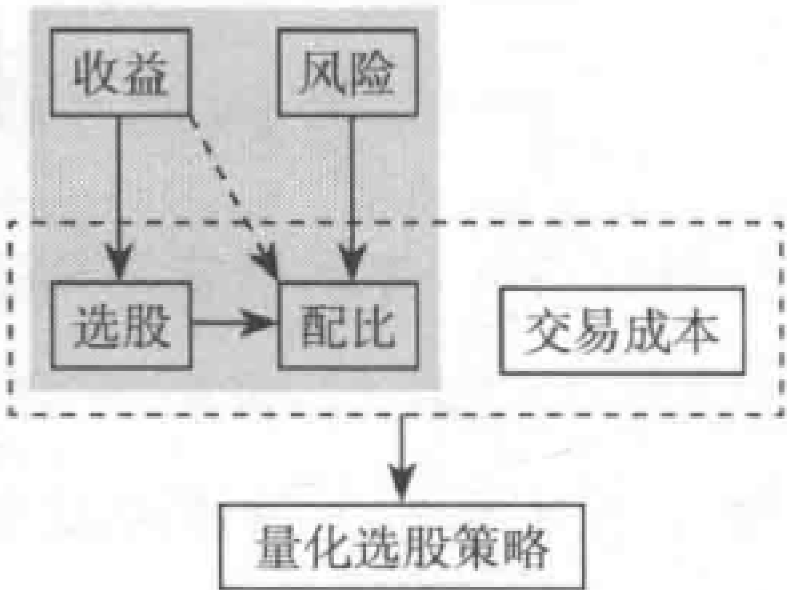


图 12-1　投资组合决策在量化选股策略研发流程中的位置

在第 1 章对量化交易策略进行简述时，我们已经提到过，马科维茨的最优投资组合理论是一个业内较为认同的量化交易策略的开端。这一经典模型又被称为均值方差理论、均值方差优化模型、现代投资组合理论等。

很明显，均值代表的是资产或投资组合的收益。至于方差，我们在第 2 章 2.2 节中也曾提到过，知道了一个投资组合中各个资产间的协方差矩阵之后，就可以根据公式计算出整个投资组合的方差，这代表了投资组合这一整体的风险水平。

方差的引入，以及多资产间协方差矩阵的引入，是最优投资组合理论最大的突破所在。在该理论问世之前，虽然也存在着“不要将所有鸡蛋放在同一个篮子里”一类的说法，但是如何具体进行分散化投资始终没有一个准确的解决方案。

而通过协方差矩阵来定量刻画资产间同涨同跌的程度，以及构造相应的模

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

型，研究者第一次能够完全定量地进行分散投资的优化决策。随着对最优投资组合理论这一类方法的深入研究以及在实践中的大量应用，协方差矩阵作为多资产情况下的风险刻画指标已经被多数人接受。

下面我们基于数学模型简单介绍一下马科维茨的最优投资组合理论。假设目前有 N 个风险资产，研究者的目标是基于这 N 个风险资产，决定投资于各个资产上的具体金额，从而最终得到的投资组合最优。

由于实际的可用资金往往因为资产价格的变动而变化，因此针对各个资产占总资产金额的比例进行决策会是一个更为便利的设定。同仓位决策的情况相似，“比例”这一工具在投资组合决策中也是具有实际作用的。

设 N 个风险资产的资金占比分别为 w_1，w_2，…，w_N，各个资金占比均不小于零。研究目标就是决策出最优的 N 维向量 $\boldsymbol{W}$，其中 $\boldsymbol{W}=(w_1, w_2, \cdots, w_N)'$。由于 N 个风险资产组成了投资组合整体，因此应该满足下式

$$\sum_{i=1}^{N} w_i = 1$$

假设构建的投资组合在某一段时间内持有，我们需要知道这段时间内 N 个风险资产的预期收益率，分别设为 r_1，r_2，…，r_N。同时，我们也需要知道这段时间内 N 个风险资产之间的预期协方差矩阵，设为

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1N} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2N} \\ \vdots & \vdots & & \vdots \\ \sigma_{N1} & \sigma_{N2} & \cdots & \sigma_{NN} \end{pmatrix}$$

式中，σ_{ij} 是资产 i 和资产 j 之间的预期协方差，σ_{ii} 代表的，实际上就是资产 i 的预期方差。

基于这些信息，可以计算得到这段时间内整个投资组合的预期收益与预期风险。由于可加性，投资组合的预期收益 R_P 实际上就是 N 个风险资产预期收益的线性组合，计算式如下

$$R_P = \sum_{i=1}^{N} w_i r_i$$

投资组合的预期风险通过预期方差来表示。但是方差不具备这种简单的可加

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

性。为了得到投资组合的整体方差 V_p，我们需要结合协方差一起进行计算，矩阵形式的计算式如下

$$V_p = W'\Sigma W$$

同时考虑投资组合的预期收益和预期风险，选取一个合适的“收益 / 风险”组合，就可以通过最优化得到最终的资产配比结果。一般而言，这一过程可以通过一个带约束的最优化问题来描述，求解这一最优化问题就可以得到资产占比 W 并应用到具体投资当中。

一个比较常见的最优化问题，是在固定好投资组合整体收益的基础上，最小化投资组合的整体风险。数学上可以描述如下

$$\min_{W} \ W'\Sigma W$$

约束条件

$$\sum_{i=1}^{N} w_i r_i = R_M,\ \sum_{i=1}^{N} w_i = 1,\ w_i \geqslant 0\ (i = 1, 2, \cdots, N)$$

其中 R_M 是需要提前确定的投资组合收益水平。

这是一个带有多个约束条件的最优化问题。作为一个二次优化问题，我们可以较为方便地得到 W 的最优解。对于这一最优问题具体的求解过程，这里就不再进行详细解释，请有兴趣的读者参阅相关资料。

另一种较为常见的最优化问题形式正好相反，是在固定好投资组合整体风险的基础上，最大化投资组合的整体收益。数学上的描述如下

$$\max_{W} \sum_{i=1}^{N} w_i r_i$$

约束条件

$$W'\Sigma W = V_M,\ \sum_{i=1}^{N} w_i = 1,\ w_i \geqslant 0\ (i = 1, 2, \cdots, N)$$

式中，V_M 是需要提前确定的投资组合方差水平。

还有一种最优化问题的形式，采用的是类似于第 9 章 9.3 节中将风险和收益相结合的方式，将投资组合的整体风险和整体收益综合为一个目标函数进行优化。例如，使用第 9 章 9.3 节曾介绍过的收益率减去方差作为优化目标，用数学

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

语言描述如下

$$\max_{W} \ W'R - \lambda W'\Sigma W$$

约束条件

$$\sum_{i=1}^{N} w_i = 1, \ w_i \geqslant 0 (i = 1, 2, \cdots, N)$$

其中 $R=(r_1, r_2, \cdots, r_N)'$，使得优化目标可以完全用矩阵形式表示。$\lambda$ 是一个用以描述投资决策者风险偏好的参数。λ 应该为正，这样目标函数就对风险给予了惩罚，符合实际情况。投资者较为厌恶风险时，对相同风险给予的惩罚力度更大，反映在参数设置上就是 λ 较大；投资者较为偏爱风险时，对相同风险给予的惩罚力度更小，反映在参数设置上就是 λ 较小。

以上就是比较常见的三种用来描述最优投资组合决策的数学模型形式。在叙述时，作者使用的是“预期收益”“预期风险”这样以未来预期为基础的描述方法，整个模型建立在对未来收益和未来风险的预测之上。如果我们基于历史样本数据来估计这段样本期内资产的收益和风险，并置于模型中进行 W 的优化，那么暗含的假设就是资产的收益和风险水平不变，估计出的历史结果就是对于未来的预期。

与此同时，我们也可以基于另外一种思路来考虑这个问题。当我们使用历史样本数据估计出这段样本期内资产的收益和风险后，就能够通过使用最优化模型，来计算出在这段时间内，资产配比 W 应该如何设定才能使得资产组合的历史表现最好。那么在未来的操作中，投资者可以沿用这一最优的配比结果。

依照这两种思路，数学上的处理过程实际上是一样的，最终做出的投资组合决策也是一样的。但是两者在内涵上有所不同，前者是基于预测做出的决策，后者则是基于最优交易系统做出的决策。读者们可以类比第 4 章 4.1 节中对两类量化择时策略的描述，来进一步理解其中的含义。

本节最后，我们简单介绍一下最优投资组合理论的变化形式。在均值方差优化的大框架下，可以通过改变具体的约束条件，使得模型能够更好地对现实交易情况进行刻画。

例如，当投资组合的构建允许卖空时，“各个资金占比均不小于零”的假设

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

就可以被放宽。我们在最优化问题当中去掉这一部分约束条件，相应的模型可以用数学描述如下

$$\max_{W} \ W'R - \lambda W'\Sigma W$$

约束条件

$$\sum_{i=1}^{N} w_i = 1$$

当然，这一个约束条件在资金使用层面上并不太符合实际。为了有效地刻画现实情况，可以设置得更复杂一点。如果在投资组合的实际构建中，不仅可以卖空，而且可以使用杠杆，那么假设做多的保证金比例是 $m_{多}$，做空的保证金比例是 $m_{空}$。此时，最优化问题变为

$$\max_{W} \ W'R - \lambda W'\Sigma W$$

约束条件

$$\sum_{w_i \geqslant 0} w_i m_{多} + \sum_{w_i < 0} (-w_i) m_{空} = 1$$

其中约束条件表示投资者在杠杆条件下，为了构建了完整的投资组合，刚好使用完所有的资金。

又比如，我们需要限制每一次重新构建投资组合时买卖资产的额度，那么可以在约束条件中加入对资产配置 W 变化的约束。设进行决策之前各个资产的资金占比分别为 $\overline{w}_1, \overline{w}_2, \cdots, \overline{w}_N$，则相应的最优化问题变为

$$\max_{W} \ W'R - \lambda W'\Sigma W$$

约束条件

$$\sum_{i=1}^{N} w_i = 1, \ \sum_{i=1}^{N} \left| w_i - \overline{w}_i \right| \leqslant U$$

式中，U 是一个提前设定的比例系数，而增加的这一个约束条件，就代表买卖资产的规模不超过总资产价值的比例 U。

通过类似这样的增加约束条件或者改变约束条件，我们就可以根据各种各样的实际需求来构建最优化问题。例如，可以限制持有资产的数量，可以限制投资组合与基准组合之间的偏离度，也可以在因子模型的基础上限制投资组合在某个

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

因子上的暴露程度，等等。

由于马科维茨的最优投资组合理论是现代金融研究的基石性理论之一，同时又对实业界的组合投资影响深远，因此不论是学术上还是实际操作上，都存在大量的关于最优投资组合理论及其内在思想的资料可供参考。而在其基础上发展出的其他相关理论，如有效前沿、资本市场线、资本资产定价模型等，更是具有丰富的内涵。有兴趣的读者应该查阅相关资料，进行进一步的学习。

12.2 实用的投资组合优化方法

虽然马科维茨的最优投资组合理论具有模型结构精炼、内涵深刻的优点，但是在实际的资产投资活动中，投资者往往更重视其思想而不是具体的模型。而且即使是纯粹的数量化决策者，往往也并不会直接套用上一节所叙述的模型来优化资产配比。这是因为，如果直接使用前面介绍的最优化模型，得到的资产配比结果常常带有不切实际的偏差。例如，如果最优化问题不带有约束条件，投资组合中许多资产的最优配置很可能就是规模非常大的空头仓位。如果最优化问题带有“禁止做空”“不设杠杆”的约束条件，某些资产的配比可能就直接被优化为零，另外一些资产则会占据过多的份额。

这种配比不均衡的最优解所导致的结果，就是模型导出的投资组合并没有得到很好的分散化，资金过度地集中于某几项资产之上。虽然使用了最优投资组合理论的数学模型，但是却失去了最优投资组合理论内在理念的支撑。一些研究者曾使用市场数据对此进行了实证分析，结果发现基于最优投资组合理论得到的投资组合，在实际表现上甚至劣于等权重的简单投资组合。

造成这一现象的原因，主要是模型的输入参数存在估计上的误差。而这种估计上的误差，可能来自第 3 章 3.2 节曾介绍过的过度拟合与欠拟合，基于历史数据的估计结果与资产的实际特性之间有差距，进而造成了样本估计与未来表现不一致。

可以看到，马科维茨的最优投资组合理论中的输入参数较多，预期收益的数据量与资产数量一致，预期协方差矩阵的数据量则是资产数量的平方。各个参数

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

中都存在或多或少的估计误差，这些估计上的误差通过整体模型混合在一起，最终会造成非常严重的偏差。

与此同时，通过最优投资组合理论计算得到的资产配比结果对输入数据的变化非常敏感。一方面，这一特性本身就不利于最优投资组合理论的实际应用；另一方面，该特性会叠加模型输入参数中估计误差，使得最优化结果与合理配比之间形成非常大的差异。

要解决马科维茨最优投资组合理论的这些缺陷与问题，研究人员提出了许多不一样的方案。这一节中作者主要介绍其中的两种方案，即改变最优投资组合理论的约束条件和 Black-Litterman 模型。

12.2.1 带约束的最优投资组合理论

在上一节中，作者已经介绍过，通过增加最优投资组合理论的约束条件或者改变其约束条件，就可以根据各种各样的实际需求来构建最优化问题，如可以限制持有资产的数量等。而增加最优化问题的约束条件，同样也是解决最优投资组合理论的缺陷的途径之一。

我们已经知道，最优投资组合理论最突出的缺陷，就是经过优化后某些资产的配比很小甚至为零，某些资产则配比过重。很直接地，就可以联想到使用约束条件来限制各个资产的最终配比，使得所有资产的占比都不会过小或过大。

首先，我们希望投资组合中的每种资产都被持有一定份额，从而保证投资组合的分散性，那么对 W 中的所有元素都设置一个下界，例如比例 L。其次，我们希望投资组合中每种资产的份额都不会过大，从而保证保证投资组合没有过度集中，那么对 W 中的所有元素都设置一个上界，例如比例 H。增加了这些约束条件之后，最优化问题就变为

$$\max_{W} W'R-\lambda W'\Sigma W$$

约束条件

$$\sum_{i=1}^{N} w_i = 1,\ L \leqslant w_i \leqslant H\ (i=1,2,\cdots,N)$$

这样的处理简单而直接，在数学设置上根本性地保证了各个资产的配置比例

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

均不越界。最后得到的优化结果符合分散化投资的核心理念，保证了对投资组合中每一项资产的均衡配置。

但是，这种处理同时也略显“粗糙”。在不改变假设条件的情况下，直接增加一组约束条件来限定最终的配置结果，会使得优化结果与原始模型的优化结果发生偏离。考虑到原始模型的优化结果本身就与“合理”的配置比例存在差异，两种差异叠加在一起，会让这种约束下的优化结果的可信程度越发受到质疑。

当然，这并不代表增加上下界约束的最优投资组合理论是一个不好的解决方案。不论是简单的模型还是复杂的模型，都存在其各自的优点与缺点，都对应着合适与不合适的应用环境。我们应该根据现实的情况与需求，来决定所使用的具体模型和方法。

例如，当我们更为关注“分散化投资”这一理念的落实时，直接在最优投资组合理论模型中增加上下界的约束就是一个相对可行的选择，甚至直接采用最简单的等权重投资组合，都可能是一个值得考虑的配比方案。

12.2.2 Black-Litterman 模型

Black-Litterman 模型是由费雪·布莱克（Fisher Black）等人提出并完善的、用于配置大类资产的投资组合决策模型。该模型可以有效地缓解资产配比不均衡的问题，同时还能够在决策中引入投资者的主观看法。

Black-Litterman 模型的建立存在一个极其重要的假设条件，即市场上各个大类资产所占的份额是均衡的，或者说由这些大类资产所组成的市场是出清的。这代表着，如果不加入投资者自身的主观判断，仅从整个市场的角度来看，当前的大类资产配比是合理的。

而 Black-Litterman 模型就从这个假设出发，基于当前的均衡情况，使用投资者的主观观点对均衡下的隐含收益率进行调整，从而导出一个被“投资者观点”影响了的新的大类资产配比结果，用以指导投资组合的构建。

模型的大致计算流程可以用图 12-2 来展示。首先观察得到市场上的大类资产配比 $W_{市场}$，并基于此计算得到均衡条件下各类资产的隐含收益率 $R_{均衡}$；然后在 $R_{均衡}$中加入投资者的主观观点，形成主观的资产收益率 $R_{调整}$；最后使用主观

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

调整后的资产收益率 $R_{调整}$，优化得到新的大类资产配比结果 $W_{调整}$。

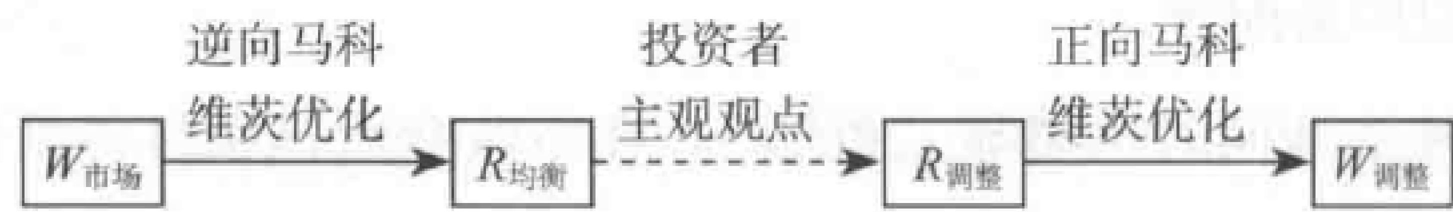


图 12-2 Black-Litterman 模型流程

图 12-2 中第一个实线箭头所代表的过程，即计算得出各类资产的隐含收益率，在数学上，就是一个逆向求解标准的马科维茨最优投资组合模型的过程。在该环境下，最优投资组合理论模型可以表示为

$$\max_{W} \ W'R_{均衡} - \lambda W'\Sigma W$$

这一最优化问题的结果是

$$W = (2\lambda\Sigma)^{-1} R_{均衡}$$

由于 Black-Litterman 模型假设市场上实际的资产配比情况就是一个均衡的结果，因此其应该与上述优化问题的结果保持一致，即有

$$W_{市场} = (2\lambda\Sigma)^{-1} R_{均衡}$$

当我们可以通过历史数据计算或其他方式得到大类资产间的协方差矩阵 $\boldsymbol{\Sigma}$ 时，就可以依据上面的式子得出均衡条件下资产的隐含收益率 $R_{均衡}$，这也是我们逆向求解标准的马科维茨最优投资组合模型所希望得到的结果

$$R_{均衡} = 2\lambda\Sigma W_{市场}$$

图 12-2 中的虚线箭头所代表的过程，则是在计算出资产的隐含收益率 $R_{均衡}$ 之后，在此基础上根据投资者的主观观点，对各类资产的收益率进行调整，从而得到新的主观资产收益率 $R_{调整}$。

需要注意的是，Black-Litterman 模型中的收益率调整，实际上调整的是资产收益率的分布。通过设定一系列参数，可以将隐含收益率的分布和主观观点对该分布的影响表示出来，两者相结合形成新的收益率分布，最后计算出新分布的期望值。这些期望值就是主观调整后资产收益率的直观表示，可以作为模型下一步的输入变量使用。具体的算式如下

$$R_{调整} = \left[(\tau\Sigma)^{-1} + \boldsymbol{P}'\boldsymbol{\Omega}^{-1}\boldsymbol{P}\right]^{-1}\left[(\tau\Sigma)^{-1} R_{均衡} + \boldsymbol{P}'\boldsymbol{\Omega}^{-1}\boldsymbol{Q}\right]$$

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

式中，$R_{调整}$是主观调整后收益率分布的期望值；τ 是一个固定的比例系数；$\boldsymbol{P}$ 是主观观点涉及资产的标注矩阵；$\boldsymbol{\Omega}$ 是用来表明观点置信度的对角矩阵；$\boldsymbol{Q}$ 是具体说明主观观点的向量。τ、$\boldsymbol{P}$、$\boldsymbol{\Omega}$、$\boldsymbol{Q}$ 均需要投资者自主进行设置。

图 12-2 中第二个实线箭头所代表的过程，则是计算出可以拿来实际使用的资产配比的过程。在数学上，就是一个普通的马科维茨最优投资组合理论模型。我们通过前面两个步骤得到了主观调整的资产收益率 $R_{调整}$，也知道各个大类资产间的协方差矩阵 $\boldsymbol{\Sigma}$，希望求出最优的资产配比 $W_{调整}$。用数学模型表示为

$$\max_{W} W'R_{调整} - \lambda W'\Sigma W$$

在这一步中，也可以根据需要加入各种约束条件，从而满足实际的投资需求。最后得到 W 的最优化结果，也就是最后用来指导投资组合构建的大类资产配比结果 $W_{调整}$。

通过观察整个 Black-Litterman 模型的运行流程可以得知，图 12-2 中的第一步和第三步其实是同一个模型正反两个方向上的求解过程。如果中间虚线箭头那一步中对收益率不做调整，那么模型得出的资产配比结果应该与市场配比保持一致。而中间步骤中对收益率的调整程度，也就决定了资产配比结果与市场配比之间的差异程度。

Black-Litterman 模型就通过这样的方法，保证了优化出的资产配置结果不会出现巨大的、不切实际的偏差。当然，既要保证配比不发生巨大偏离，又要保证投资者的主观观点对配比结果产生一定影响，就需要合适的资产收益率调整程度和参数设置作为基础。好在关于 Black-Litterman 模型的研究资料非常之多，感兴趣的读者可以参阅相关的文献进一步学习，本书限于主旨不对此进行展开。

最后需要重点强调的是，“市场上各类资产所占的份额是均衡的”这一假设，对于 Black-Litterman 模型非常重要。离开了这一假设，对资产的市场配比进行调整的做法就失去了逻辑基础。

这也是 Black-Litterman 模型一般用于对大类资产进行配置的原因，大类资产相对而言覆盖面更宽泛，因此若干大类资产之间就会更好地体现出均衡的状态。一些研究将 Black-Litterman 模型用于股票的行业配置，作者对此持保留意

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

见。特别是在中国，本书写作之时，股票发行主要还是核准制，股市的自我选择能力较弱，市场的实际情况与均衡假设是否差距过大，是值得考虑的问题。至于在个股配置等更为微观的层面上使用 Black-Litterman 模型，就越发需要仔细论证了。

这也是作者一直在书中阐述的观念，在使用一个策略或模型之前，深入的理解策略模型是非常必要的。明确策略模型的内在含义，搞清楚其具体的适用范围，才能正确地使用策略模型。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

第13章

优化的股票配置策略

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

13.1 多因子风险模型

上一章介绍了进行投资组合决策的具体模型和方法。在本章中，作者将会在第 8 章 8.1 节所介绍的多因子选股策略的基础上，进一步对策略所选取的股票进行配比决策，形成一个投资组合优化的完整股票投资策略。

在 12 章 12.2 节中，我们阐述了传统的马科维茨最优投资组合理论所具有的一些缺陷，并试着给出了两个解决方案，即改变标准最优投资组合理论的约束条件和 Black-Litterman 模型。

可以看到，改变标准最优投资组合理论的约束条件虽然略显"粗糙"，但是其实是一个相对而言适用范围较广的处理方法。而 Black-Litterman 模型则由于假设条件的限制，更适合于大类资产的配置。

由于本章所涉及的配比对象是股票而非大类资产，因此 Black-Litterman 模型就是一个不太适用的决策方法，不再加以讨论。这一节的主要内容是额外介绍一个适用于股票组合配置的处理方法。该方法中的一个重要组成部分是多因子风险模型，其可以有效地解决最优投资组合理论中待估计参数过多、估计误差过大的问题。

多因子风险模型在名称上与多因子选股模型有一定的相似之处。实际上，这两个模型渊源颇深，在"股票收益受到特征因子的影响或驱动"的核心理念上是一致的。其区别在于，多因子选股模型的目标是预测收益率，而多因子风险模型的目标在于估计股票之间的协方差矩阵。根据由这两个模型分别得到的收益率和协方差矩阵，配合马科维茨的最优投资组合理论或其他改进形式，就可以优化得到最终的股票配比结果。这一过程如图 13-1 所示。

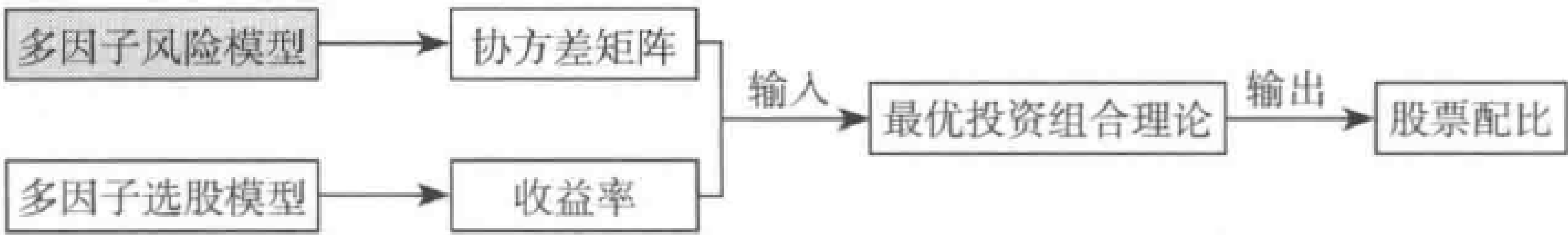


图 13-1 加入多因子风险模型的投资组合决策过程

多因子风险模型的构建基础与多因子选股模型一样，都是股票的收益与多个特征因子之间的线性关系。假定已经选出了 K 个有效的风险因子，分别用 f_1 到

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

f_K 标记，然后针对每一只股票，用这 K 个因子来表示相对应的收益率，如下式所示

$$r_t = \sum_{k=1}^{K} b_{k,t} f_{k,t-1} + u_t$$

式中，r_t 是时刻 t 的股票收益率；$f_{k,\,t-1}$ 即为 $t-1$ 时刻下第 k 个因子的大小；$b_{k,\,t}$ 是回归式中要估计的系数；u_t 是确定参数后回归所得的残差。

可以将等式右边的前半部分看作风险因子可以解释的收益，后半部分即残差则是不能解释的特质收益。对于多因子风险模型而言，有一个非常重要的假设条件，那就是可以被解释的收益与特质收益之间不存在相关性，且不同股票之间的特质收益也不存在相关性。

有了这个假设之后，我们就可以先针对风险因子部分计算出各只股票之间由于因子特征而形成的协方差矩阵，再针对残差部分计算出各只股票的特质方差，两者简单相加就是各只股票之间的协方差矩阵。

如果可以获得 T 个时间段的因子数据以及相应的下一期股票收益率数据，那么对于上面的线性回归式而言，一共可以进行 T 次估计，即 $t=1$，…，T 时刻下的估计。得到的参数 $b_{k,\,t}$ 可以组成一个 $K \times T$ 的矩阵。基于该矩阵，可以计算出 K 个因子系数之间的协方差矩阵，记为 $\boldsymbol{B}$。

再设 T 时刻下各只股票的因子大小分别为 $f_{k,\,T}$，可以形成一个 $K \times T$ 的矩阵，记为 $\boldsymbol{F}$，其中 N 是 T 时刻的股票只数。㊀那么在 T+1 时刻下，这 N 只股票之间由于因子特征而形成的协方差矩阵可以表示为 $\boldsymbol{FBF'}$。

然后，针对这 N 只股票，计算其在 $t=1$，…，T 时刻下的回归残差，通过这 N 个残差序列统计得到各只股票的特质方差。设计一个 $N \times N$ 的对角矩阵，记为 $\boldsymbol{\Delta}$，其对角线上的对应项为各只股票的特质方差，其他项均为零。

综合上述结论，可以预测在 T+1 时刻下，这 N 只股票之间的协方差矩阵为

$$\boldsymbol{\Sigma} = \boldsymbol{FBF'} + \boldsymbol{\Delta}$$

从模型设置上来讲，多因子风险模型与多因子选股模型并不要求使用完全一

㊀ 实际运作中，需要考虑到数据缺失、新股发行等问题所导致的无效股票剔除，因此 N 往往要小于某一时刻下所有股票的只数。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

样的因子。原因在于，虽然影响股票收益的因子往往会影响股票收益的风险，但是反之则不然。一些特征因子并不显著地影响股票收益，但是却会使得股票收益的波动变化，或者使得股票之间由于该因子的联动性导致股票间的相关性发生改变，这些因子都有理由加入到多因子风险模型当中。

常见的不影响股票收益但是应该列入风险因子行列的因子包括 Beta[⊖]、行业分类等。Beta 是根据资本资产定价模型，股票和市场组合同涨同跌的程度，这种同涨同跌的性质，也会相应地形成股票之间同涨同跌的现象。行业分类更容易理解，相同行业的股票往往由于经营项目的相似性，因此在涨跌上具有一定的趋同，这就会影响到股票之间的相关程度，进而影响到股票组合的风险特征。

一般而言，多因子选股模型中使用到的因子都应该加入多因子风险模型。因为因子既然能直接影响到股票收益，那么或多或少都会对收益的风险产生影响。与此同时，在多因子风险模型中加入与多因子选股模型相同的因子，还有一个额外的效用。在多因子选股模型重点配置具有某些因子特性的股票时，多因子风险模型又会因为相同因子产生的风险而降低这些股票的配置，这有利于股票组合的合理分散化。

这里介绍的多因子风险模型，实际上是最为基本和简单的设置形式。关于模型的进一步细化，以及各个风险因子的选取，读者可以参考其他的相关资料。[⊜]

13.2 投资组合优化的多因子策略

上一节介绍了适用于股票组合风险估计的多因子风险模型，图 13-1 和相关内容也给出了如何使用多因子风险模型和多因子选股模型来共同构造一个股票投资组合的过程。在本节中，作者将基于第 8 章 8.1 节所介绍的推进的多因子选股策略，进一步计算股票组合的风险，在选股之后得出对应股票的最优配比，从而构造出一个更为完整的股票配置策略案例。

⊖ 法玛的论文验证了 Beta 与预期收益无关。但是这种无关性，在 Beta 可以被视作风险的前提下，又能够用来制作 Smart Beta 策略，此为题外话。

⊜ 如 BARRA 的 USE3 等。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

由于增加了多因子风险模型，对比第 8 章 8.1 节来看，使用的数据有所扩展。所选股票的范围依然是在深圳证券交易所和上海证券交易所进行交易的所有 A 股。月度数据则包括之前用到的月末收盘时股票的复权后月度收益率和股票的月度流通市值，时间从 1991 年 1 月开始直到 2015 年 6 月底结束。

然后，进一步增加同期的股票月度 Beta 数据，这部分数据仅用于计算协方差矩阵的多因子风险模型当中。加上之前的市值因子和反转因子，这一节中的多因子风险模型一共涉及三个风险因子，如图 13-2 所示。流通市值同样进行了对数化处理，从而更好地反映出月度收益率和流通市值之间的线性关系。

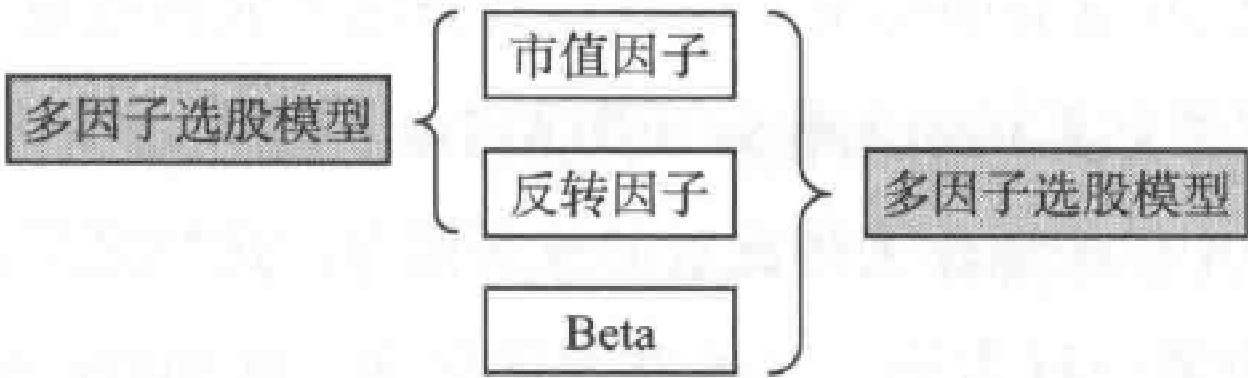


图 13-2 多因子选股模型和多因子风险模型的因子选取

大的推进分析框架依然维持第 8 章 8.1 节中的设定，只是在各个推进步骤中都相应地加入了多因子风险模型和最优投资组合计算的处理。图 13-3 展示了多因子策略下股票配置策略的推进分析在这段时间的样本下是如何进行的。该图和图 8-1 是相同的，也就是说每一行中优化、检验的路径与第 8 章 8.1 节是一样的。

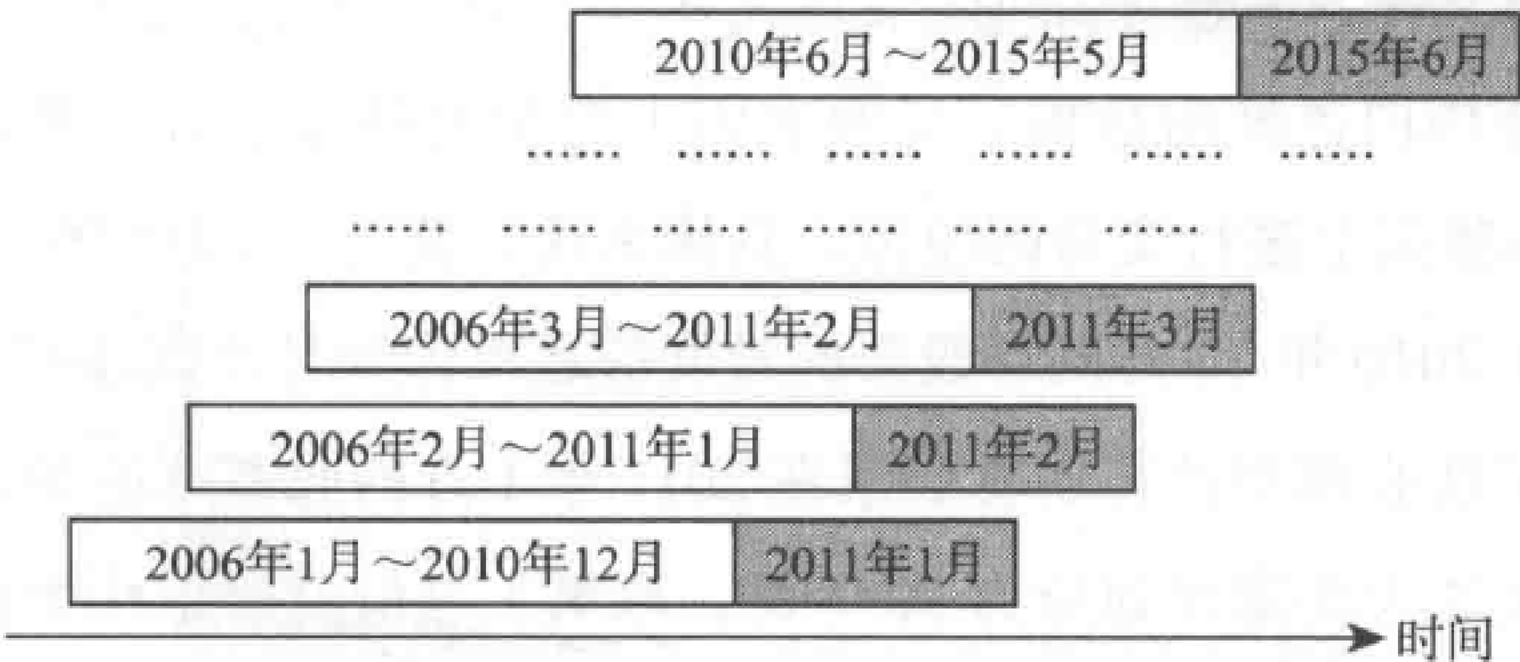


图 13-3 多因子策略下的最优投资组合的推进分析流程

具体的区别就在于每一行中，白色优化框都增加了使用多因子风险模型计算协方差矩阵的步骤，灰色检验框中的股票则不再是等权重配置，而是根据所有股

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

票的收益率预测值和所有股票之间的预期协方差矩阵，构造最优投资组合理论模型，优化得出各只股票的最优占比，进而进行投资组合的配置。

在第 8 章 8.1 节中，我们在预测出各只股票在灰色框检验期中的收益率之后，按照所有股票收益预期大小排序划分成了 3 等份、10 等份、30 等份共三种情况，然后各自取第一个等份，即收益预期最高的等份中的股票，按照等权重进行组合。第一等份组合的收益率序列就是策略的模拟交易收益情况。

我们知道，待配置的股票数量越少，最优投资组合理论模型实际需要的样本量就越小。而正如我们前面提到过的，中国市场的数据量本身比较有限，因此这一节中我们只研究 30 等分的情况，也就是划分出的股票数量最少的情况。首先通过多因子选股模型计算得出的收益预期大小，将所有股票分成 30 个等份；然后选择第一个等份中的股票作为待配置的股票组合，进行之后的优化配置。

以推进分析的第一行为例，白色框为优化期，从 2006 年 1 月到 2010 年 12 月一共 60 个月。首先复制第 8 章 8.1 节的处理，根据这 60 个月的月度数据，结合 2005 年 12 月的备用数据，得到 60 组数据，从而进行多因子选股模型的回归估计。

其次，增加股票的月度 Beta 数据，同样得到 60 组数据，进行多因子风险模型的回归估计。基于得到的 60 组系数估计值，计算出三个因子系数之间的协方差矩阵，即上一节提到过的矩阵 $\boldsymbol{B}$，同时计算出各只股票在回归之后的残差的方差，用来形成上一节中的特质风险对角阵 $\boldsymbol{\Delta}$。

再次，使用白色框的数据完成两个因子模型的估计之后，则将回归式用到灰色框的样本数据上进行交易的模拟。具体来说，就是在 2010 年 12 月收盘的时间点上，基于 2010 年 12 月所有股票的月度流通市值和月度收益率数据，使用估计出的多因子选股模型计算所有股票在 2011 年 1 月的收益率的预测值。将所有股票按收益预期大小排序划分为 30 等份，取第 1 等份的股票作为待配置的股票组合。

再在 2010 年 12 月收盘的时间点上，基于这些待配置股票的月度流通市值、月度收益率和月度 Beta 数据，使用前面已经得到的矩阵 $\boldsymbol{B}$，以及由待配置股票的残差方差形成的对角阵 $\boldsymbol{\Delta}$，就可以计算出待配置股票之间的协方差矩阵，具体

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

计算公式参见上一节。

那么，在 2010 年 12 月收盘这个时间点上，就可以选择出待配置的股票组合，并得到这些股票的预测收益率，以及这些股票的预测协方差矩阵。接下来，就可以使用最优投资组合理论来对这些股票进行优化配比。在本节的实例中，统一设置风险偏好参数 $\lambda=1$，那么最优化模型可以描述如下

$$\max_{W} \ W'R-W'\Sigma W$$

约束条件

$$\sum_{i=1}^{N} w_i = 1,\ w_i \geqslant 0(i=1, 2,\cdots, N)$$

通过这些步骤得出的股票选择与配置，就是在第一行灰色框，即 2011 年 1 月间持有的股票组合；再结合这些股票在 2011 年 1 月的月度收益率，就可以计算出整个股票组合的收益情况。

至此，第一行的优化与检验的一整套步骤就完成了。在处理完这一行的操作之后，将白色框和灰色框都分别向后推进一个月，重复相同的排序和等分过程，直到最后一行为止。

首先来看一下模拟交易过程中，股票组合内各只股票的配置情况。由于检验期一共 54 个月中每个月的股票组合都不相同，全部说明显得冗余，因此这里以 2011 年 1 月、2 月、3 月的情况为例进行展示。图 13-4 中的柱状图从左到右，分别表示 2011 年 1 月的股票组合中，占比高到占比低的股票各自的资金占比比例。

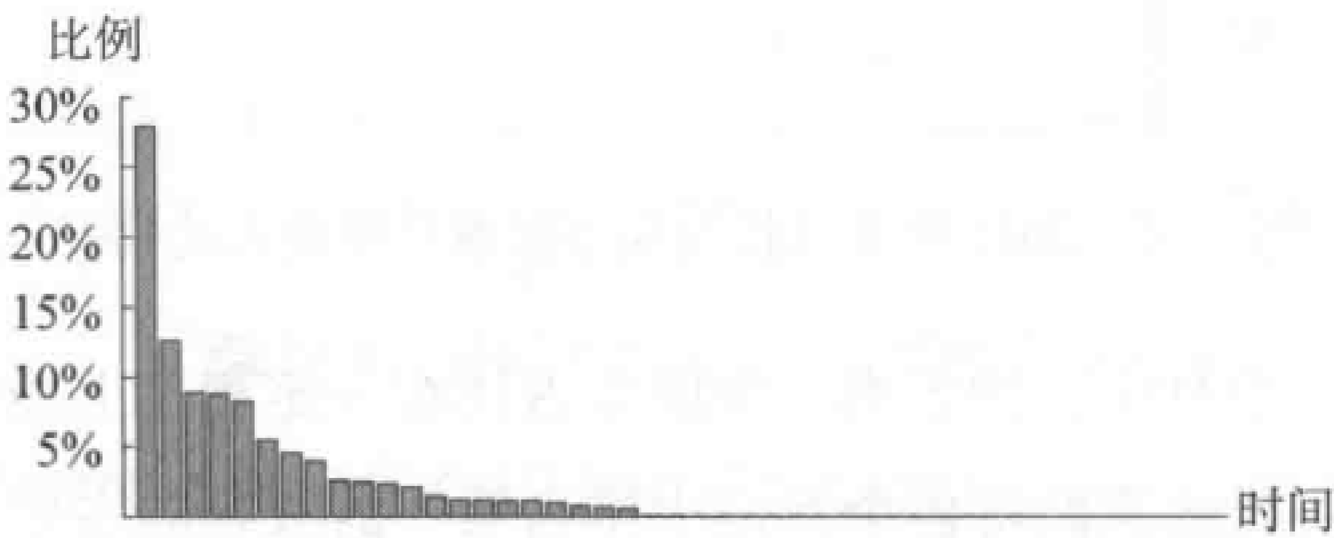


图 13-4　2011 年 1 月检验期下的股票组合占比情况

由图 13-4 我们可以看到，2011 年 1 月占比最高的股票，其资金比例为 27.96%。资金占比第二的股票降为 12.66%，第三的则是 8.91%。占比前三名的

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

股票，一共占据了总资金量的 49.52%，接近一半。

图 13-5 展示了 2011 年 2 月检验期下的股票组合占比情况。相比起 2011 年 1 月的配比结果，这一个月中的配置更为分散化。排名第一的股票的资金占比为 10.20%，占比第二的股票与占比第一的股票相差无几，为 10.10%，第三名的占比则为 8.37%。同样计算前几名的合计占比，发现前六名的股票一共占据 49.81% 的资金量。

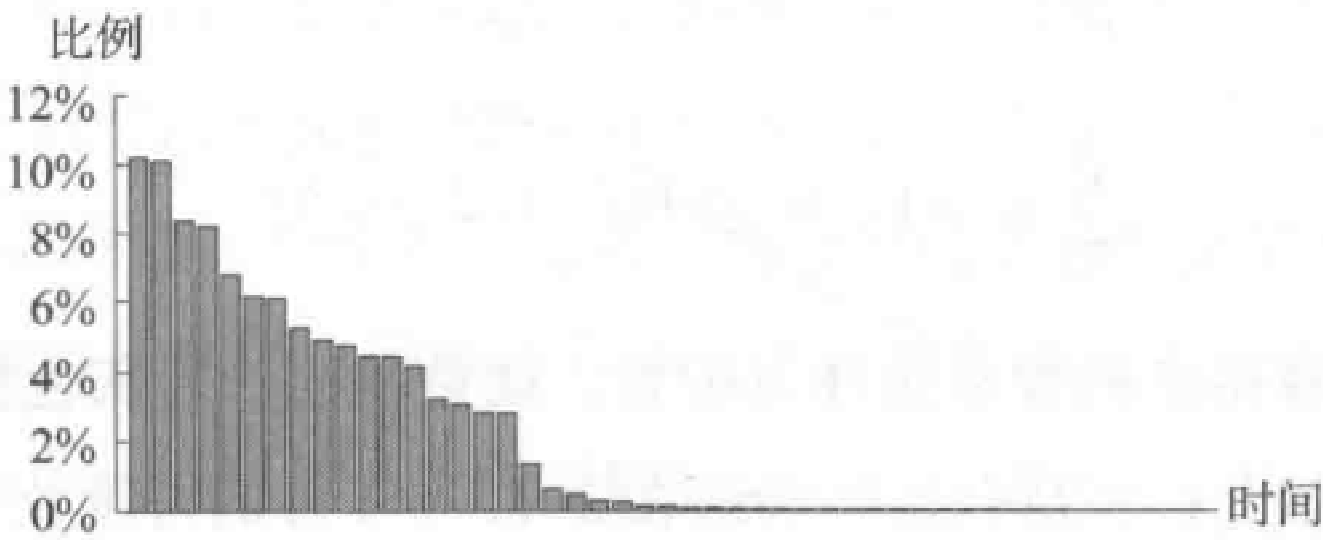


图 13-5　2011 年 2 月检验期下的股票组合占比情况

图 13-6 展示的是 2011 年 3 月检验期下的股票组合占比情况。与前两个月相差较大的是，其资金占比最高的股票，占据了总资金量的 75.95%。第二名的股票则直接降至 3.84%，是第一名的近 1/20。在图 13-6 中我们也能很清楚地看到，整个股票组合的配置过分集中于第一只股票，分散程度不佳。

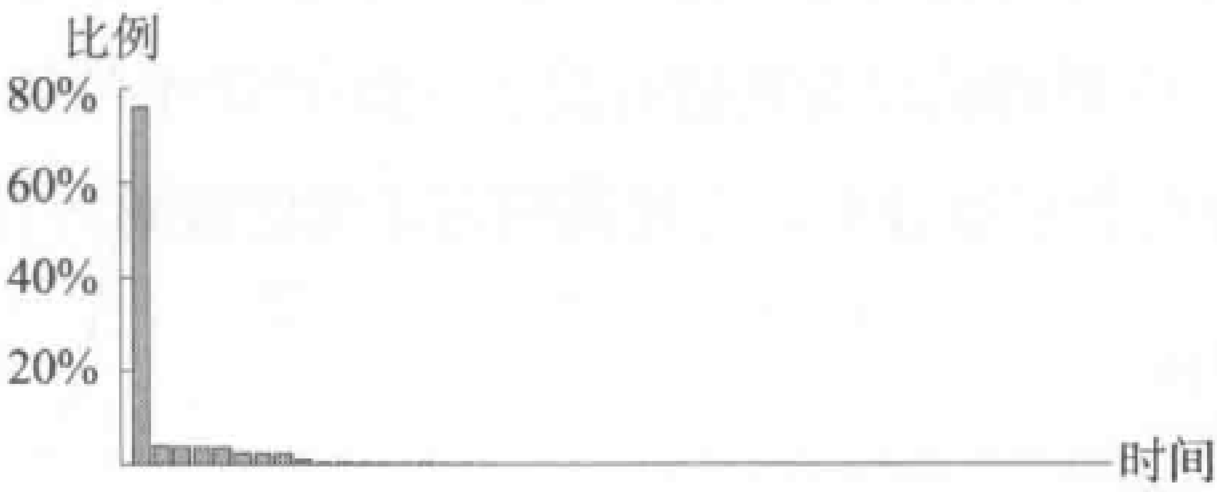


图 13-6　2011 年 3 月检验期下的股票组合占比情况

检验期的每一个月中，排名第一的股票占比在一定程度上能够反映股票组合的整体配比情况。因此我们将各个月份中排名第一的股票的占比合并为一个时间序列，长度为 54，将这个时间序列制成一个柱状图，即图 13-7。

从图 13-7 中我们可以看到，四年半的检验期中，排名第一的股票占比并没有明显的趋势特征。股票配置较为极端的有两个月份，一个是前面提到的

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

2011 年 3 月，另一个则是 2013 年 1 月，该月份中单个股票的最高占比达到了 69.94%。整体来看，月度最高占比的平均值是 21.48%。对比这个平均值可知，前面展示的 2011 年的前三个月，实际上是具有一定代表性的三个月。

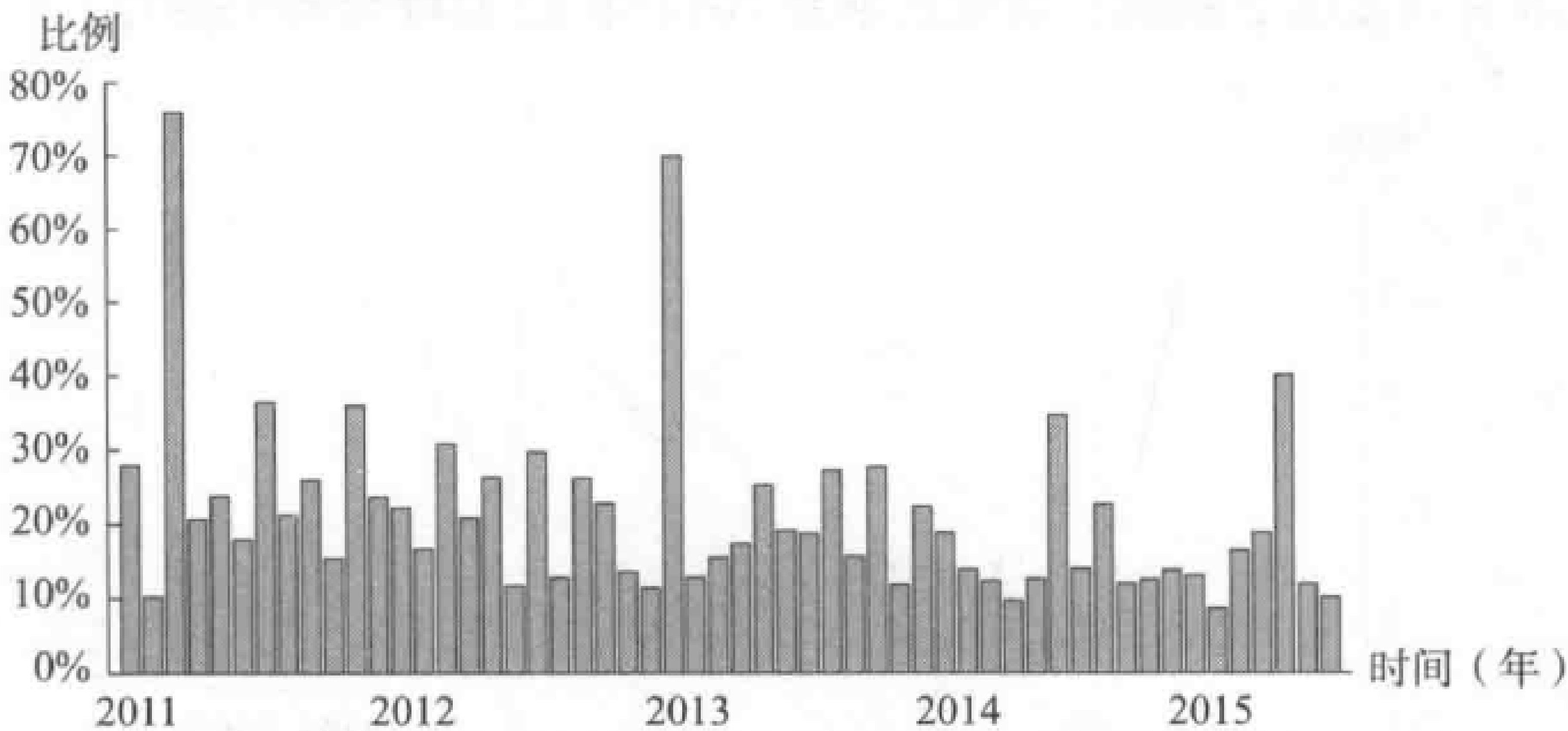


图 13-7 检验期内逐月投资组合中最高配置的股票占比情况

针对检验期的 54 个月，我们可以优化得出每个月的具体股票配置情况，并计算出相应的股票组合收益率。那么，根据这一长度为 54 的收益率序列，就可以计算出因子模型下的股票配置策略在灰色框模拟交易的净值走势。

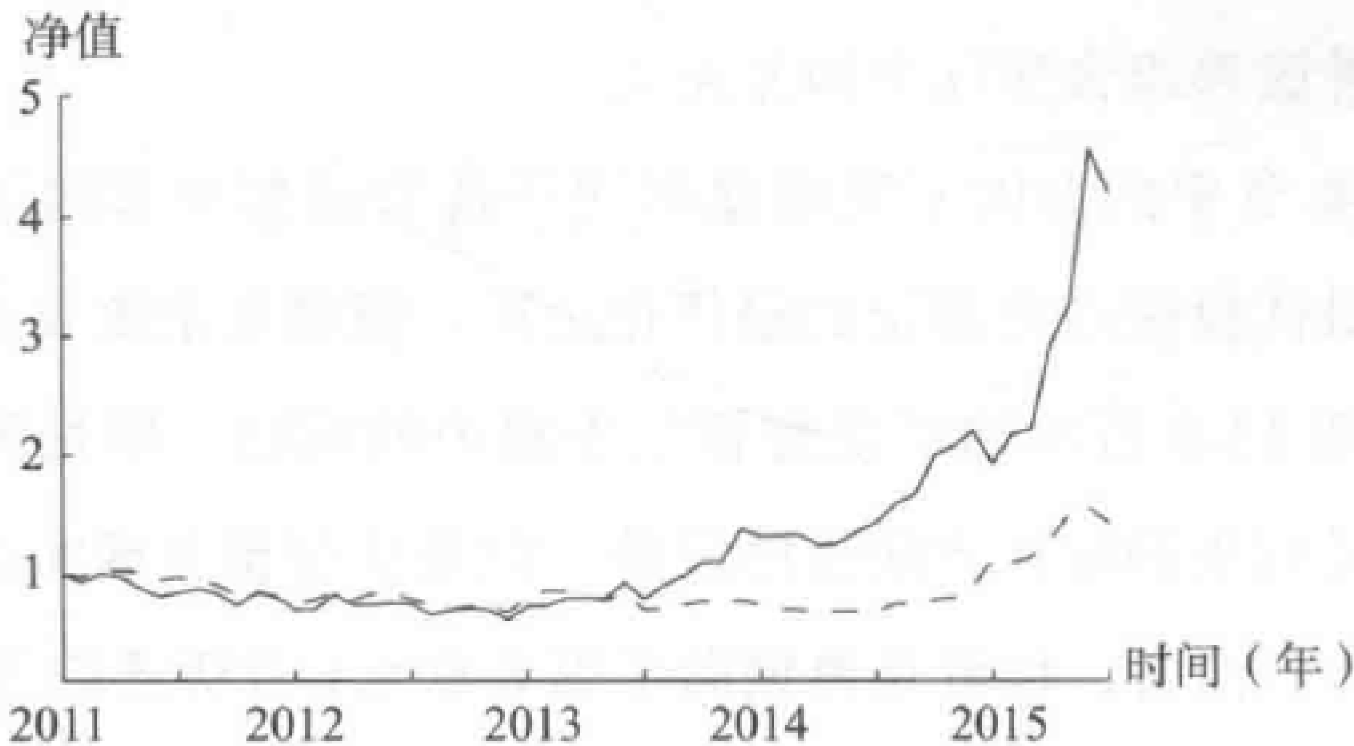


图 13-8 多因子策略下的股票组合净值走势

图 13-8 中的实线展示了使用多因子策略进行股票配置的情况下，以 1 为初始值的策略净值走势，虚线则是用来进行对比的同期沪深 300 指数走势。策略净值在 2015 年 6 月收盘时达到 4.2003，四年半的策略收益率为 320.03%，折算下来的年化收益率为 37.56%。

图 13-9 则是逐月地将多因子策略下的股票组合收益率减去沪深 300 指数的

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

收益率后，所得到的收益率序列形成的净值走势。净值从 2010 年 12 月收盘的 1 开始，在 2015 年 6 月底达到 2.6823，四年半的收益率为 168.23%，折算下来的年化收益率为 24.52%。四年半之间的最大回撤率为 41.98%，回撤的开始点为 2014 年 10 月收盘的 2.6494，结束点则是 2014 年 12 月收盘的 1.5372。

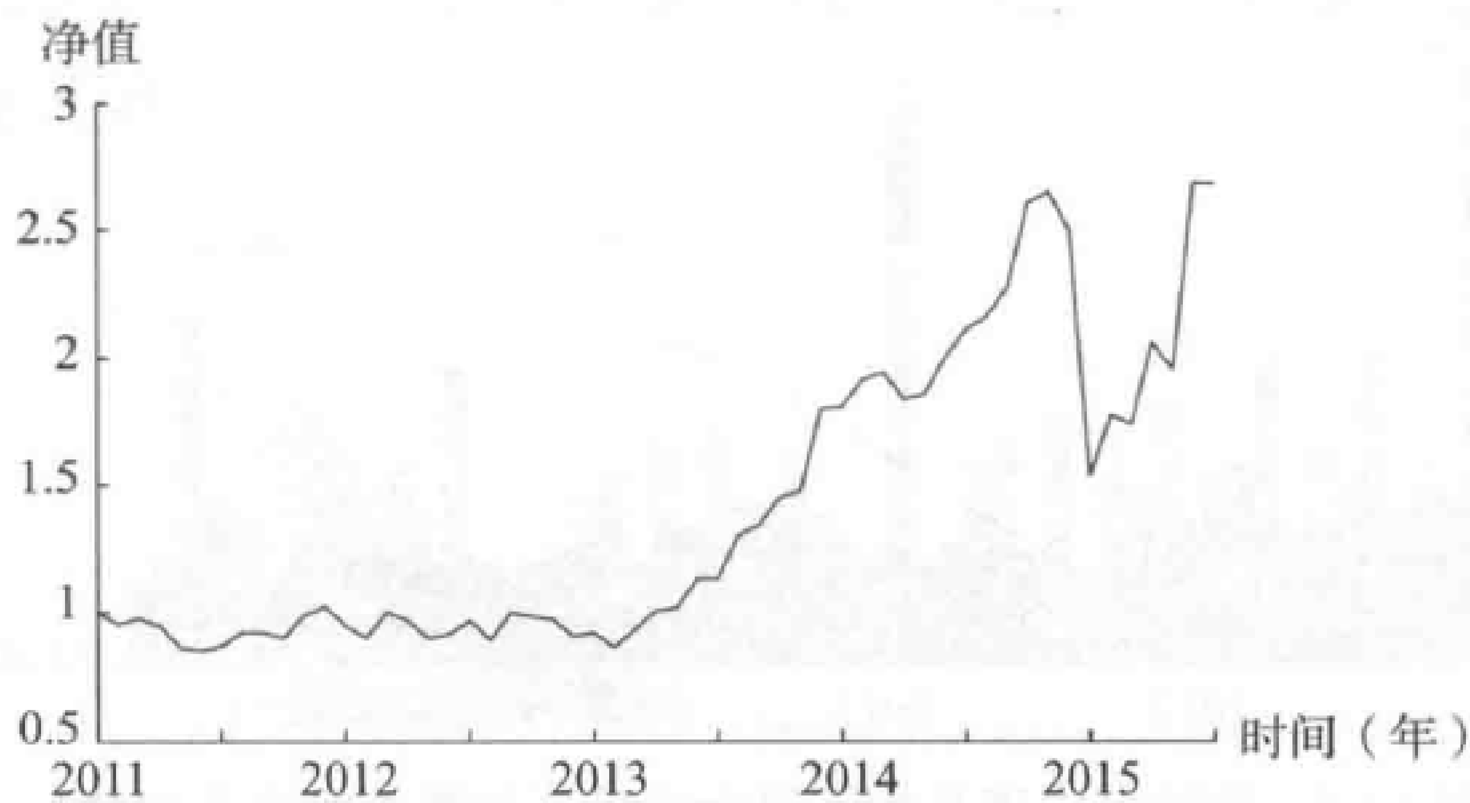


图 13-9　多因子策略下的股票组合与沪深 300 收益之差的净值走势

与第 8 章 8.1 节中的等权重组合对比来看，不论是股票组合的净值走势，还是股票组合与沪深 300 指数收益之差的净值走势，都表现得非常相似。多因子策略下配置优化的股票组合具有略小的回撤程度，但是收益情况也略低。计算收益和回撤之比，两种股票组合实际上相差无几。

但是考虑到本节中的多因子策略是多因子选股模型和多因子风险模型的组合，同时增加了最优投资组合理论的最优化过程，模型复杂度要高出不少，这是一大劣势。诸如图 13-6 所示的股票配置过于集中的问题，则是策略实际体现出的另一个劣势。等权重的配比方法虽然简单，但是从逻辑上确实达成了投资分散化的目的。两相比较而言，作者是更偏向于第 8 章 8.1 节所述的等权重策略的。

而要让多因子风险模型和最优投资组合理论更好地发挥作用，我们需要更多有效的风险因子、更好的模型拟合，很可能需要更多的数据以及一定程度上的人工干预。要让看似更为精准的理论落实到实际交易层面上，是一个复杂的实践工程。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

第14章

交易成本

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

14.1 交易成本估计

交易成本是指在交易策略的运行过程中，除去资产价格的变动外，需要额外承担的净值变化，而这种变化主要是损耗性的、负面的，因此称为“成本”。在量化交易策略的研究中，一般可以将交易成本分为两个大的方面，即显性交易成本和隐性交易成本。其中显性交易成本主要是交易税费和买卖价差，隐性交易成本则主要包括滑点和冲击成本，如图 14-1 所示。

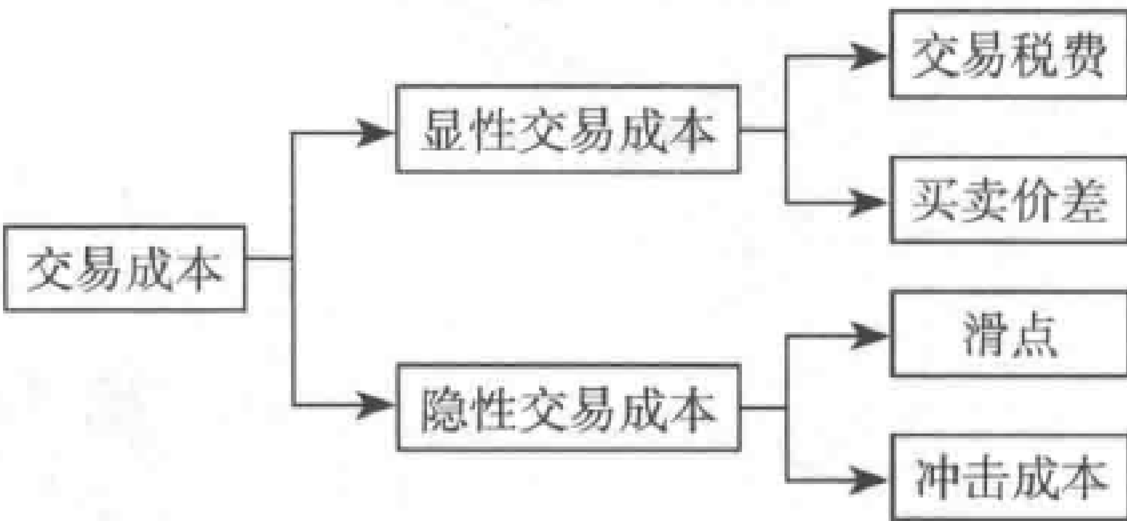


图 14-1 交易成本的分类

交易税费是这几项中最容易理解的，这实际上是当我们在进行交易的时候，向我们提供交易这项服务的机构所收取的服务费用。例如交易所为我们提供了交易的场所，就会收取交易经手费；证券公司为我们提供了通向交易所的通道和交易平台，就会收取交易佣金；政府为我们提供交易在法律上的保障，因此收取印花税，等等。交易者支付这些税费，其实是在为交易过程中所获得的服务付费。

相比起其他种类的交易成本，交易税费在形式上更为明确。特别是二级市场，交易相关的服务其实都是标准化的，因此针对服务所收取的费用也相应地明确而标准。当然，交易者可能会由于交易量较大等原因而获得交易佣金等方面的优惠，不过优惠之后的具体费用也应该是明确的。由此可知，在量化交易策略的研发过程中，交易税费可以被较为方便而准确地计算出来。

买卖价差是显性交易成本中比较特殊的一种，其是指在同一个时间点上，买入和卖出同一个资产的交易价格之间的差异。在报价驱动的做市商环境下，做市商主动给出买卖报价，并赚取两者之间的价差作为提供交易服务的报酬。而在指令驱动市场中，买卖差价由未成交的限价单决定，并不是某个特定对象收取的服务费。市场流动性与买卖价差有直接的相关性，当市场的流动性较差时，买卖价差往往较高，因此有学者将可观测的买卖差价直接作为市场流动性的度量指标。

滑点属于隐性交易成本，指的是下单的价格点位和最后成交的价格点位之间的差距，这种差距产生的原因在于下单与真正成交之间存在一定的时间差。由于

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

市场上的资产价格随时间而不断变化，因此在“下单”与“成交”这两个时间点之间就会形成价格上的差异。我们可以用一个公式来简单说明：

滑点 = 下单与成交之间的时间差 × 资产价格变动速度

当然，该公式对资产价格的变化形式做了简化，将其视为一个匀速运动。可以看到，滑点的变化，主要源于等式右边的两项，即时间差和变动速度。由于在这段时间差内，资产价格的变动方向并不确定，因此滑点对净值而言不一定就是损耗性的，如果在时间差内资产价格向上变化，滑点甚至会造成资金的额外增值。不过不论资产价格向上向下，滑点都会导致净值不确定性的增加，这是其负面效应所在。

针对滑点的研究，可以依据该公式分别落实在右边的两项之上。在资产价格变动速度相同的条件下，下单与成交之间的时间差越长，滑点情况就越严重，净值不确定性也就越大。而在下单与成交之间时差固定的情况下，资产价格的变动速度越快，净值的变化也就越大，滑点越为严重。

为了控制滑点的程度，交易者会努力缩短从下单到实际成交的时间。这些努力包括：用程序化代替人工下单从而节省人做判断的时间，提升硬件条件从而降低从下单到交易所接收的延迟等。而交易所接收后到实际成交之间的时间长短，则往往与市场流动性密切相关，需要研究者提前预估。

资产价格的变动速度也是在研究滑点时需要进行预估的内容。特别的，当有重大新闻公布，或者发生一些影响较大的突发事件时，市场往往会出现反应，资产价格变动速度短时间内大幅度升高，交易者应特别注意这种情况。

冲击成本是指交易本身对资产价格产生影响，进而所导致的损失。例如，当我们要买入一个资产时，买入行为增加了需求，从而拉升了资产的价格，因此需要在更高的价位上买入资产，从而侵蚀了交易的整体收益。而当我们要卖出一个资产时，卖出行为增加了供给，从而降低了资产的价格，我们只能以更低的价格卖出资产，因此侵蚀了交易的整体收益。

可以看到，冲击成本实际上是交易行为在市场上的自我反馈，其所造成的结果基本都是损耗性的。所执行的交易量越大，交易对价格的冲击程度就越强，所形成的冲击成本就越高。当然，冲击成本与市场容量、市场流动性也有直接关

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

系。市场本身的体量越大、市场的流动性越好，其所能容纳的资金量就越大，同等交易量所造成的价格冲击也就越小。

作为隐性交易成本，滑点和冲击成本都没有确定的形式。它们是交易行为在市场环境中自然形成的成本，耗散于市场之中。相较于显性交易成本，隐性交易成本的估算往往更为复杂和困难。

表 14-1 对这几种交易成本进行了归纳。需要注意的是，估算难度并没有一定的难易差别，买卖价差、滑点、冲击成本等在学术上都有大量的研究成果，估计模型也各有不同。准确说来，还是估计的精准度要求越高，估算难度就越大。

表 14-1　各类交易成本总结

交易成本种类	估算难度	与流动性关系	可忽略的情况
交易税费	容易	无关	无
买卖价差	一般	相关	无
滑点	较难	相关	交易频率低
冲击成本	较难	相关	资金量小

对于一个完整的量化交易策略研发流程而言，形式更为明确、估算相对方便的交易税费，还是应该尽量加以计算并用于研究的。这些某种意义上的服务费用是发生交易时确定要付出的成本，没有理由在研究时完全忽略。买卖价差则需要视情况而定，例如在做市商仅通过买卖价差收取服务费用的情况下，对买卖价差的估计就显得很有必要。

对于交易频率较低的交易策略来说，下单与成交之间的时间差显得较为短暂，在资产价格变动速度没有巨大变化的前提下，所产生的滑点相较于更长时间上的资产价格变化，在一定程度上可以被忽略。

对于资金量较小的情况而言，由于交易量对资产价格的影响不大，冲击成本也就能够被忽略掉。当然，这需要同时判断市场的容量和流动性，这里所说的资金量小其实是一个相对的说法。

上述内容就是我们对于交易成本的粗浅归纳。一般情况下，研究量化交易策略时考虑到这几个种类的交易成本就足够了。虽然在经济学层面，交易成本有着更为深刻的内涵，常常包含机会成本等组成部分。㊀但是在量化交易策略的研究

㊀ 参见罗纳德·科斯（Ronald Coase）的交易成本理论等。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

中，所谓的交易成本更为现实，更贴近我们平常所理解的“成本”。如果要考虑量化交易策略的机会成本，那就需要考虑这笔资金如果不用于这一策略而用于其他策略，可能获得的收益。除了少数配置型策略的情况，作者看不到其作用所在。

介绍完几种主要的交易成本之后，再来看交易成本计算在量化交易策略研发流程中的位置，如图 14-2 所示。在本书介绍的基本流程中，对交易成本的考量放在买卖和仓位决策做出之后，实际上独立于这两者，目的在于使得回溯测试更接近于实际交易环境。考虑了交易成本之后，量化交易策略的回溯测试业绩一般会降低。我们在判断是否实际使用该策略时，理应针对这一业绩进行判断，而不是针对那个未考虑成本的、较之实际更好的业绩。

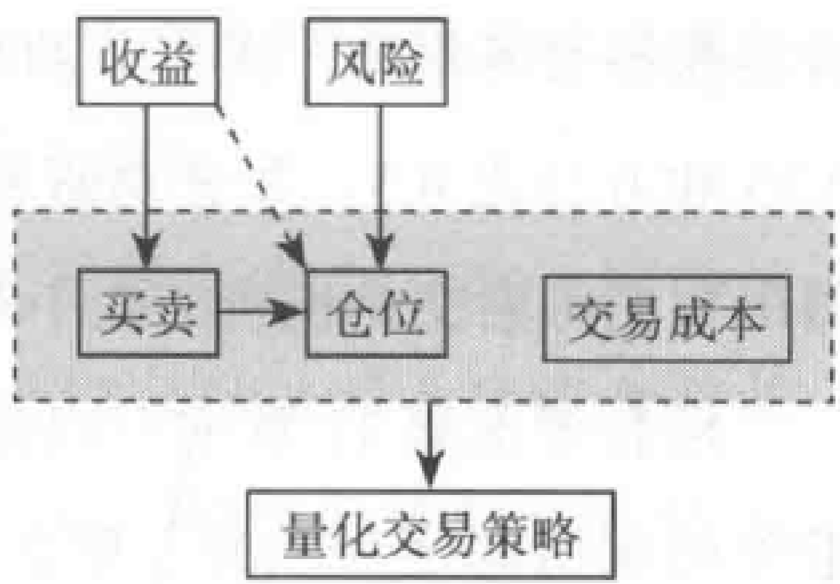


图 14-2　交易成本计算在量化交易策略研发流程中的位置

需要注意的是，虽然这个策略研发流程的核心似乎是买卖和仓位的决策过程，但是我们不应该轻视交易成本计算在整个流程中的作用。买卖、仓位所构成的交易策略在实际操作中的盈利能力并不是百分之百确定的，但是交易策略在实际操作中是一定会形成交易成本的。实际上，更重视这种百分之百会出现的负面影响，才是对自己的研究负责任的表现。

作者在下面两节会分别针对第 11 章曾讨论过的择时策略和第 13 章曾讨论过的选股策略，考虑交易成本的影响，得到相对而言更有参考价值的回溯测试结果。由于本书的实际案例并没有关于资金量等背景的假设，同时也是出于简化实例的目的，下两节的交易成本分析仅考虑最简单的交易税费。如果读者对滑点、冲击成本等的估算有兴趣，作者建议自行参考相关的学术研究。

14.2　考虑交易成本的择时策略

我们的择时策略的交易资产是股指期货，而股指期货收取的交易税费比较简单。通过查阅中国金融期货交易所对于沪深 300 股指期货产品的说明可知，该产品的手续费为万分之 0.25，双边收取。提供经纪业务的期货公司会在交易所手续

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

费的基础之上，额外收取一定比例的手续费，作为自己的服务费用。

不同的期货公司会根据公司状况、客户资金量等情况额外收取高低不等的手续费。一些期货公司也会选择不收取额外的手续费，仅通过沉淀资金的利息等收益获得服务报酬。在本节的案例中，我们考虑两种手续费收取情况，即万分之0.25和万分之0.5，并分别研究这两种情况下的择时策略。这样的设置，也可以帮助我们考察交易税费的大小对策略表现的影响。

综合考虑第11章中不同标准差阈值下的策略表现情况，这一节中我们将标准差的阈值固定为0.02。那么，我们将要研究的就是四种情况：万分之0.25手续费时的均线趋势策略；万分之0.5手续费时的均线趋势策略；万分之0.25手续费时的自回归策略；万分之0.5手续费时的自回归策略。

14.2.1 万分之0.25手续费时的均线趋势策略

图14-3展示了与第11章11.1节设置相同的均线趋势策略，在万分之0.25手续费的条件下，模拟交易过程以1为初始值的净值走势情况。在2012年年底、2013年年底和2014年年底，策略的净值分别变为0.8183、0.6218和1.2689。净值整体走势与第11章11.1节中的策略净值走势基本一致，只是因为交易成本的原因使得表现有所下降。

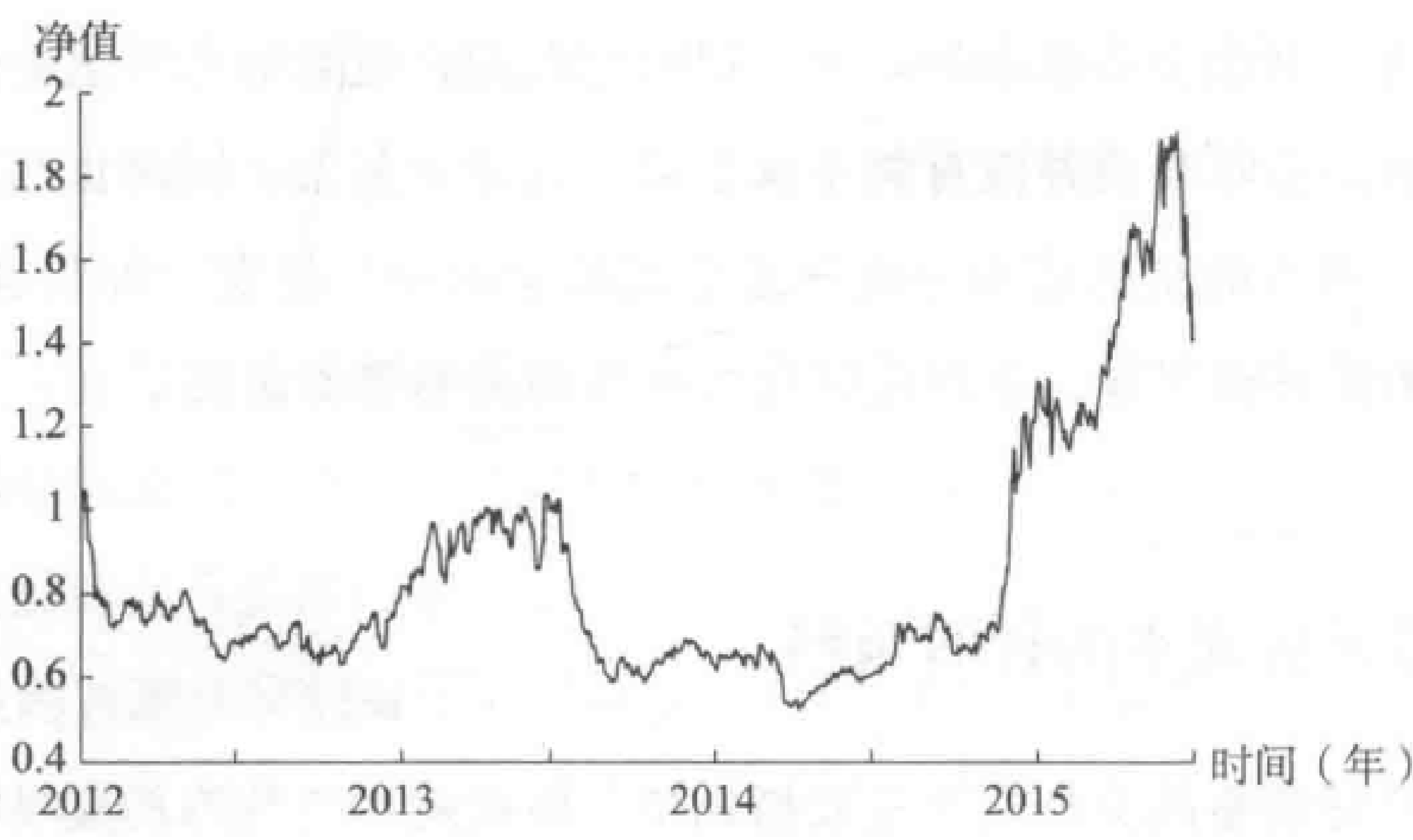


图14-3 万分之0.25手续费下的均线趋势策略净值走势

净值同样从2014年年末开始呈现出一个上升的态势，在2015年6月12日

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

达到最高点 1.9058，最后一段时间有一个较为明显的下跌，到 2015 年 6 月 30 日收盘时的策略净值为 1.4036。

基于整个交易过程来看，三年半的策略收益率为 40.36%，年化收益率为 10.17%。对比来看，不考虑交易成本时的整体收益率是 49.54%，年化收益率为 12.18%。考虑交易税费之后的年化收益率下降幅度为 2.01%。

在整个三年半的模拟交易中，策略净值的最大回撤率为 50.05%，回撤的发生从 2012 年 1 月 5 日收盘的 1.0446 开始，一直到 2014 年 4 月 9 日收盘时的 0.5218 结束，整个回撤过程一共经历 543 个交易日。这一最大回撤期与第 11 章 11.1 节中的回撤期不同，但是与第 7 章 7.1 节中的回撤期一致。

经过计算，均线趋势策略在整个交易过程中平均的年化最大回撤率为 30.06%。相较于不考虑交易成本时的 29.71%，年化最大回撤的程度有细微的变大。

14.2.2 万分之 0.5 手续费时的均线趋势策略

图 14-4 展示了与第 11 章 11.1 节设置相同的均线趋势策略，在万分之 0.5 手续费的条件下，模拟交易过程以 1 为初始值的净值走势情况。在 2012 年年底、2013 年年底和 2014 年年底，策略的净值分别变为 0.8032、0.5996 和 1.1998。净值表现因为交易成本增加的原因进一步下降。

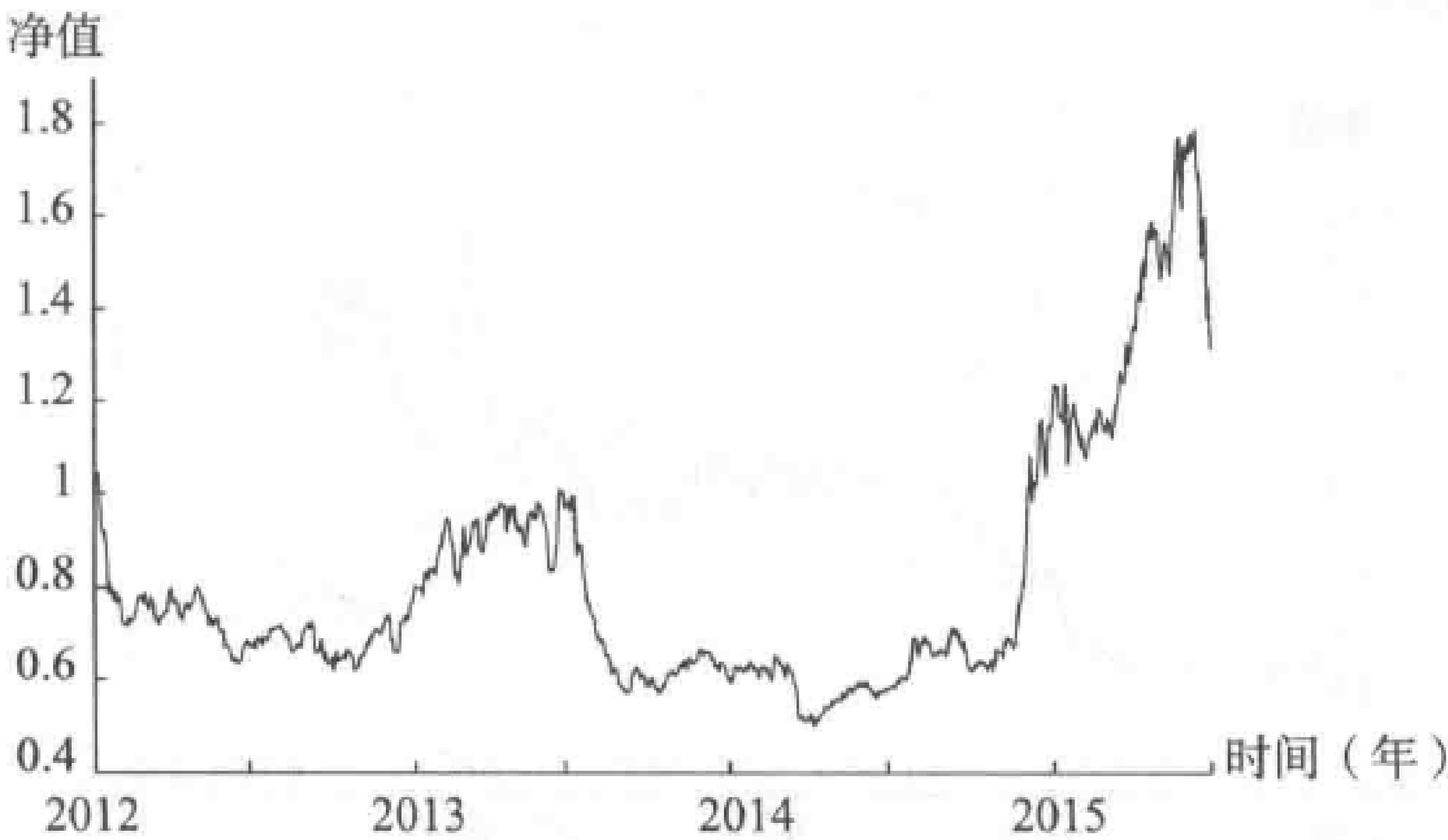


图 14-4 万分之 0.5 手续费下的均线趋势策略净值走势

同样的，净值从 2014 年末开始呈现出一个上升的态势，在 2015 年 6 月 12 日达到最高点 1.7898，最后一段时间有一个较为明显的下跌，到 2015 年 6 月 30

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

日收盘时的策略净值为 1.3174。

基于整个交易过程来看，三年半的策略收益率为 31.74%，年化收益率为 8.20%。对比来看，不考虑交易成本时的年化收益率为 12.18%，手续费为万分之 0.25 时的年化收益率为 10.17%。当手续费由万分之 0.25 提升到万分之 0.5 后，年化收益率又下降了 1.97%。

在整个三年半的模拟交易中，策略净值的最大回撤率为 52.04%，回撤的发生从 2012 年 1 月 5 日收盘的 1.0444 开始，一直到 2014 年 4 月 9 日收盘时的 0.5009 结束，整个回撤过程一共经历 543 个交易日。最大回撤期与手续费为万分之 0.25 时的情况保持一致，回撤幅度略有增加。

经过计算，均线趋势策略在整个交易过程中平均的年化最大回撤率为 30.42%。相较于万分之 0.25 手续费时的 30.06%，年化最大回撤的程度也有细微的加大。

14.2.3 万分之 0.25 手续费时的自回归策略

图 14-5 展示了与第 11 章 11.2 节设置相同的自回归策略，在万分之 0.25 手续费的条件下，模拟交易过程以 1 为初始值的净值走势情况。在 2012 年年底、2013 年年底和 2014 年年底，策略的净值分别变为 1.5599、2.0538 和 3.1287。与第 11 章 11.2 节中的策略净值走势相比，这里的策略净值也因为交易成本在原有基础上有所下降。

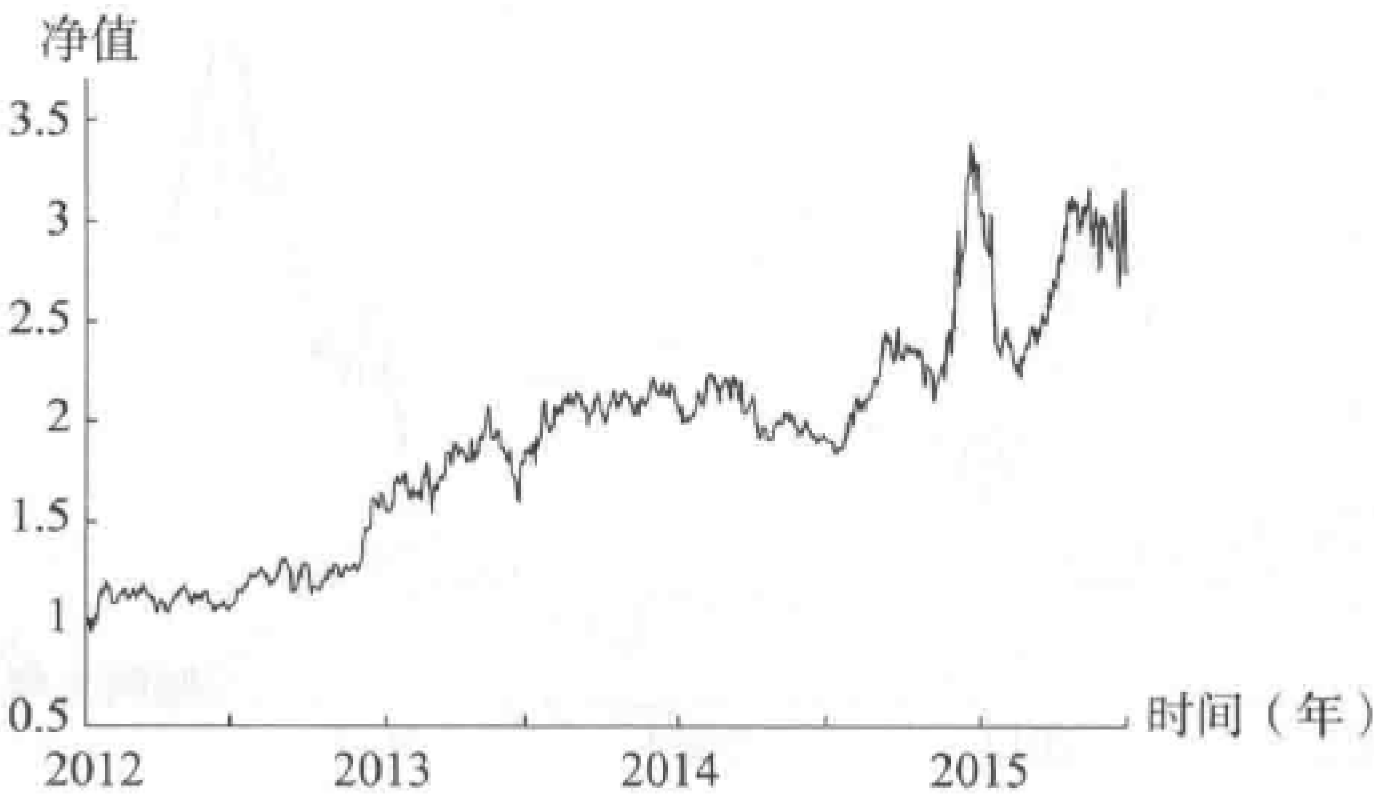


图 14-5 万分之 0.25 手续费下的自回归策略净值走势

净值同样在 2014 年 12 月 22 日达到最高值，数值为 3.3819，在最高点附近呈现出一个波动相对较小的倒 V 字形状。样本期期末，即 2015 年 6 月 30 日收

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

盘时的策略净值为 2.7387。

基于整个交易过程来看，三年半的策略收益率为 173.87%，年化收益率为 33.36%。对比来看，不考虑交易成本时的整体收益率是 191.96%，年化收益率为 35.82%，考虑交易税费之后的年化收益率下降幅度为 2.46%。虽然对比相同手续费设置下的均线趋势策略，自回归策略受交易成本影响的降幅略大，但是考虑到该策略的收益水平本身更高，实际上手续费对收益的侵蚀比例是较小的。

在整个三年半的模拟交易中，策略净值的最大回撤率为 34.27%，回撤的发生从 2014 年 12 月 22 日收盘的 3.3819 开始，一直到 2015 年 2 月 25 日收盘时的 2.2231 结束，整个回撤过程一共经历 40 个交易日。这一最大回撤期与不考虑交易成本时的情况一致，只不过回撤幅度略有加大。

经过计算，自回归策略在整个交易过程中平均的年化最大回撤率为 21.74%。相较于不考虑交易成本时的 21.48%，年化最大回撤的程度有细微的变大。

14.2.4 万分之 0.5 手续费时的自回归策略

图 14-6 展示了与第 11 章 11.2 节设置相同的自回归策略，在万分之 0.5 手续费的条件下，模拟交易过程以 1 为初始值的净值走势情况。在 2012 年年底、2013 年年底和 2014 年年底，策略的净值分别变为 1.5309、1.9796 和 2.9563。

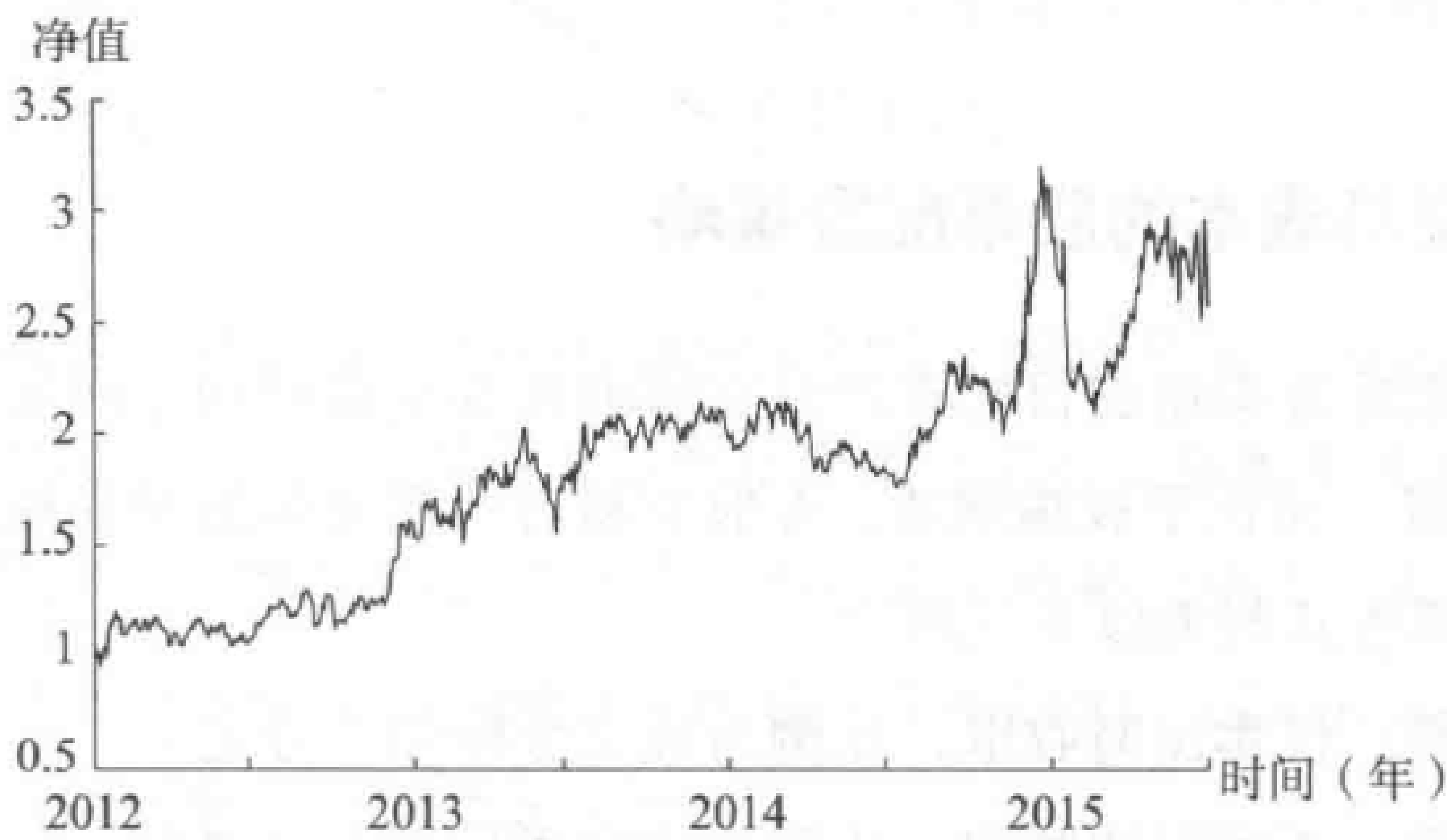


图 14-6 万分之 0.5 手续费下的自回归策略净值走势

可以看到，净值表现因为交易成本增加的原因进一步下降。在 2014 年 12 月 22 日净值达到最高点，为 3.1973，并同样呈现出一个波动相对较小的倒 V 字形

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

状。样本期期末，即 2015 年 6 月 30 日收盘时的策略净值为 2.5690。

从整个交易过程来看，三年半的策略收益率为 156.90%，年化收益率为 30.94%。对比起来，不考虑交易成本时的年化收益率为 35.82%，手续费为万分之 0.25 时的年化收益率为 33.36%。由于手续费由万分之 0.25 提升到万分之 0.5，年化收益率又下降了 2.42%。

在整个三年半的模拟交易中，策略净值的最大回撤率为 34.44%，回撤的发生从 2014 年 12 月 22 日收盘的 3.1973 开始，一直到 2015 年 2 月 25 日收盘时的 2.0960 结束，整个回撤过程一共经历 40 个交易日。最大回撤期依然保持一致，回撤幅度又略有增加。

经过计算，自回归策略在整个交易过程中平均的年化最大回撤率为 22.00%。相较于万分之 0.25 手续费时的 21.74%，年化最大回撤的程度也有细微的加大。

可以发现，在自回归策略中考虑交易成本的影响之后，交易成本对策略业绩的侵蚀相较于均线趋势策略的情况较小，这主要是因为本书案例中的自回归策略的收益情况更好。

但是我们有必要在这里重申。回溯测试展示出的收益情况并不一定能在实盘中重现，而且由于过度拟合的原因，实盘业绩往往要比回溯测试的结果更差。但是交易成本在实际操作中是一定会产生的，因此更加认真地对待交易成本的负面影响才是合乎逻辑的研究态度。

14.3 考虑交易成本的股票配置策略

我们的股票配置策略的目标资产是在深圳证券交易所和上海证券交易所进行交易的所有 A 股。相比于股指期货，A 股交易者所要支付的交易税费的项目稍多一些，这里大概将其归为以下三项。

（1）印花税：仅卖出时收取，比例为成交金额的千分之一。

（2）过户费：买卖双边收取，比例为成交面值的万分之三。仅对在上海证券交易所交易的股票收取。

（3）佣金：买卖双边收取。最高不超过成交金额的千分之三，最低 5 元起，单笔交易佣金不满 5 元按 5 元收取。具体收取水平由提供经济业务的证券公司决定。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

除此之外，A 股的交易税费还包括证券监管费和证券交易经手费。这两项费用合在一起俗称交易规费，现在一般都由证券公司代收，即合并到佣金中统一收取。因此在上面三项交易税费的归纳中，没有将这两种费用单独列出。

在本节的例子中，我们将证券公司收取的佣金设定为成交金额的万分之三。那么针对印花税、过户费和佣金三项费用，都有了明确的收取规则，我们也就能够准确地将交易成本考虑到交易策略的回溯测试中去。

下面作者将基于与第 13 章 13.2 节设置相同的股票组合策略，使用上述交易税费设定，重新模拟交易过程，并研究交易策略的业绩表现。

需要注意的是，该交易策略是在每个月月末的交易时点确定新的股票配置比例。由于交易成本的加入，因此在建仓、换股过程中会遭受资金的损失。只有确定了建仓、换股后新的资金总量，才能确定新的股票配置中各只股票具体的资金份额，从而用于实际投资交易。

确定月末交易后新的资金总量的方法有两种：一种较为简单，就是假设换仓操作没有交易成本，那么各只股票的交易量就是交易前后股票配置比例之间的差额，依此交易量计算相应的交易税费，就可以通过减去该交易税费得到新的资金总量。

另一种则是在资金利用率 100% 的条件下，假设交易前的资金总量为 1，交易后的资金总量为 x，根据这两个资金总量分别计算交易前后各只股票的金额大小，金额之差就是各自的交易量。然后依此交易量计算出相应的交易税费，其应该是 x 的一个函数 $f(x)$。那么应该存在如下等式：

$$1-x=f(x)$$

通过求解该等式，可以得到具体的 x 的值，也就是新的资金总量。而将每个月的收益率乘以 x，就可以得到考虑交易成本之后的交易策略收益率。

本节的案例计算采用第二种方法。不过由于两种方法之间差异一般较小，选择实际使用中更为简单的第一种方法也是完全可以的。这里只是提供一个数学模型解决实际问题的例子，供读者参考。

那么，针对检验期的 54 个月，我们考虑交易成本并计算出相应的股票组合收益率。进而，根据这一长度为 54 的收益率序列，就可以计算出股票配置策略在模拟交易下以 1 为初始值的净值走势情况。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

图 14-7 中的实线展示了相应的策略净值走势，虚线则是用来进行对比的同期沪深 300 指数走势。策略净值在 2015 年 6 月收盘时达到 3.8718，四年半的策略收益率为 287.18%，折算下来的年化收益率为 35.10%。对比而言，不考虑交易成本时的年化收益率为 37.56%。考虑交易税费之后的年化收益率下降幅度为 2.46%。

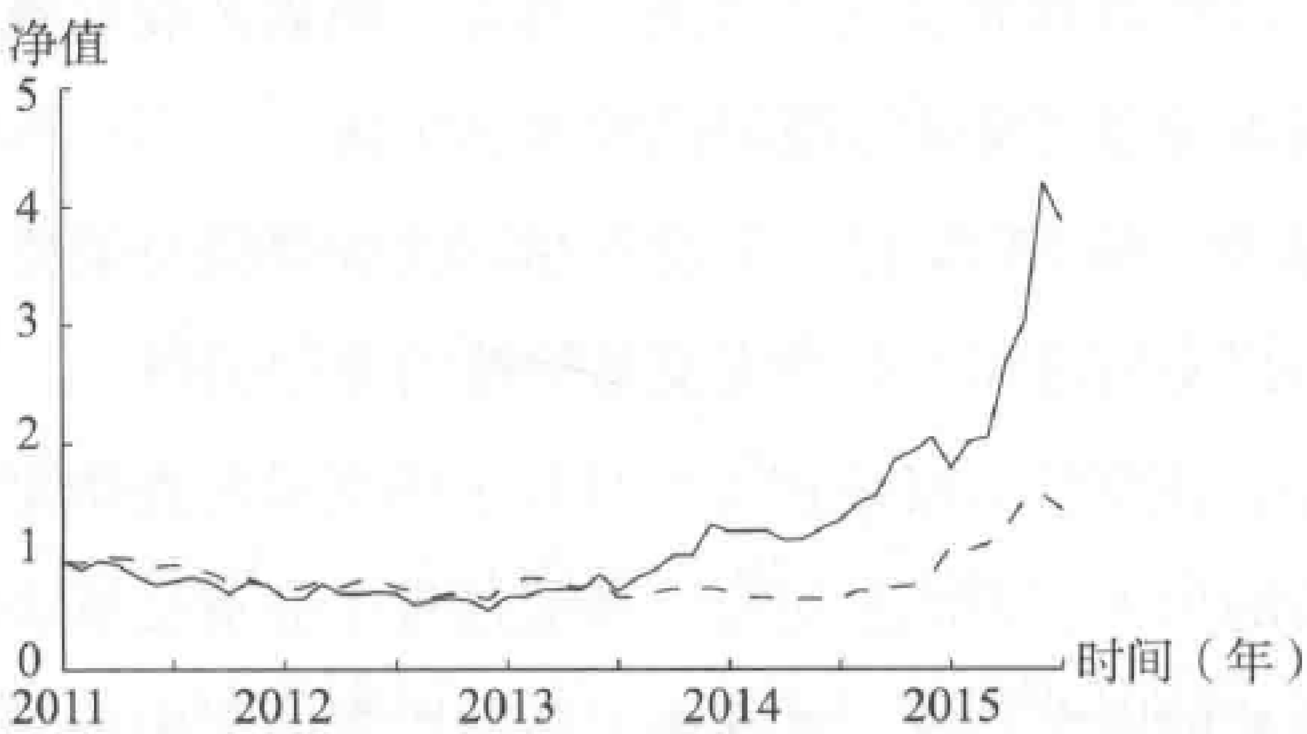


图 14-7 考虑交易成本后的股票组合净值走势

图 14-8 则是逐月地将考虑交易成本的股票组合收益率减去沪深 300 指数的收益率后，所得到的收益率序列形成的净值走势。净值从 2010 年 12 月收盘的 1 开始，在 2015 年 6 月底达到 2.4700，四年半的收益率为 147.00%，折算下来的年化收益率为 22.25%。不考虑交易成本时，相应的年化收益率为 24.52%。考虑交易税费之后的年化收益率下降幅度为 2.27%。

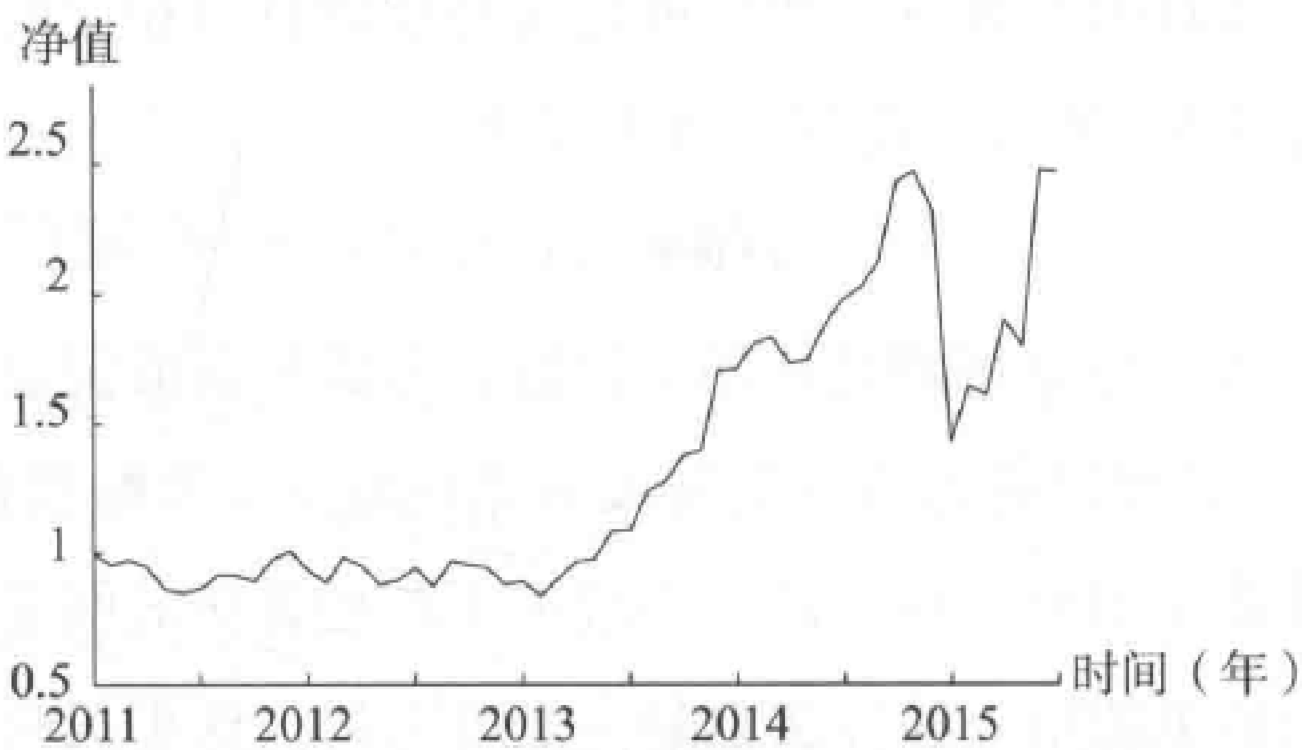


图 14-8 考虑交易成本后的股票组合与沪深 300 收益之差的净值走势

四年半之间的最大回撤率为 42.20%，回撤的开始点为 2014 年 10 月收盘的 2.4697，结束点则是 2014 年 12 月收盘的 1.4275。最大回撤期与不考虑交易成本的情况保持一致，回撤幅度稍有增加。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

第15章

策略评价

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

15.1 策略评价体系

到上一章为止，我们针对所选取的择时策略案例和选股策略案例，已经完成了第 2 章中图 2-1 所示的一个完整的研究流程。驱动整个研究进程的是：交易策略的收益、风险，以及将收益和风险相结合的度量指标。

而在完成了整个研究流程之后，研究人员往往需要将工作中所涉及的关于交易策略的一些情况进行汇总，进而对策略形成一个总体的评价。这种做法既方便了研究人员对自己的交易策略形成一个整体上的把握，同时也能够提供给其他人，帮助他们对交易策略进行认知和评判。因此，我们可以将策略评价看作量化交易策略研发的一项额外但是有必要的收尾工作。

从实际情况出发，策略评价可以用于两种情况：一是针对交易策略的回溯测试，对其模拟交易的业绩进行评价；二是针对交易策略的实盘结果，根据其实际的净值走势进行评价。由于作者较为推荐回溯测试尽量贴近于实盘交易，因此在作者自己的研究和交易体系中，这两种情况下的策略评价方式也是基本一致的。只不过对于这两种情况下策略评价结果的态度应该不尽相同，毕竟实盘操作是落实的结果，而回溯测试往往会由于过度拟合等问题出现失真。

而对于评价体系的构建，实际上可以沿用策略研究当中所涉及的几个方面，即收益、风险、风险调整后收益，再结合其他一些有必要加入评价体系的描述性内容，从而完成对量化交易策略的整体评价。

15.1.1 收益

评价交易策略收益情况的常用指标较少，但是都非常重要。一个指标是整个交易期间的整体收益率，它是用来描述交易策略在所研究时间段内取得收益的硬性指标。另外一个常见的指标是年化收益率，通过将整体收益率换算为一年的收益率，研究者就可以更为方便地对不同的策略和产品进行横向对比。

最大盈利率和最大亏损率也是两个可以加入策略评价的收益类指标，其分别代表策略从交易期初期开始，在整个交易期内达到的最高的收益水平和最低的收益水平。实际上，这两者对应的就是整个交易期内策略净值的最高点和最低点。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

需要注意的是，在第 9 章讨论风险时，我们曾经将最大亏损率归为风险度量指标一类，认为其可以刻画策略的下侧风险。而在实际应用中，这些偏向于对量化交易策略的实际表现进行描述的指标，往往存在多个方面的特质。这里放在收益或者风险类别当中都是可以的，对于使用者来说更重要的还是正确理解其含义。

15.1.2 风险

在第 9 章的前两节中，作者已经对一些有价值的风险指标进行了论述。这里可以沿用之前的思路，将风险指标分为两类：一类是偏向于描述净值走势的；另一类是偏向于对收益率序列做统计的。对比而言，在策略评价中前者更为直观一些。

偏向于描述净值走势的，包括第 9 章曾经提到过的最大回撤率和年化最大回撤率。这两个指标都是将策略损失的可能性作为风险的主要来源，进而刻画策略净值下跌绝对值的风险度量方式。

而在实际的策略交易过程中，损失持续的时间长度往往也是一个很重要的不利因素。因此在策略评价报告中，也可以考虑加入对最大回撤具体发生时间的展示，以及对年化最大回撤时长的计算等。

同样根据第 9 章的内容，我们选取若干个偏向于对收益率序列做统计的指标，包括方差、半方差、标准差和风险价值 VaR。其中的标准差可以再进行一次年化计算，给出年化标准差以方便对比。

15.1.3 收益和风险的结合

可以参考第 9 章 9.3 节的内容，选取合适的风险调整后收益指标，来综合评价交易策略的收益和风险。由于收益的评价指标比较单一，因此具体指标间的区别主要就是风险指标的区别，即分为偏向于描述净值走势的和偏向于对收益率序列做统计的。

偏向于描述净值走势的指标，可以选取如下两个：

年化收益率 / 最大回撤

年化收益率 / 年化最大回撤

而偏向于对收益率序列做统计的指标，可以选取夏普比率、索提诺比率，以

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

及当比较基准明确存在时的信息比率。为了方便策略之间的横向对比，可以将这几个比率进行年化转换。

15.1.4 其他

除了收益、风险、风险调整后收益以外，还有其他一些评价内容可以用来有效地对交易策略进行描述，我们将其都归在这一类之中。其中比较基本的，包括交易策略进行模拟交易或者实盘交易的时间范围，还有非常有价值的净值走势图。不同于前面所提到的指标，净值走势图可以更加完整地描述出交易策略在整个交易期间的表现情况。我们可以将前面的指标看作压缩后的信息，方便研究人员更为直接地针对某一个特征进行观察。而净值走势图则展示了策略交易表现得更为完整的信息，可以从不同角度去评判研究。

在第 10 章 10.1 节介绍凯利公式时，我们曾经提到过胜率和盈亏比的概念。作为交易环节的概括性指标，这两者也常常被纳入策略评价当中。相应的，还可以展示用来计算胜率的盈利日数、亏损日数，以及用来计算盈亏比的盈利日平均盈利、亏损日平均亏损。这是日度数据的情况，如果使用月度数据则换为盈利月数等即可。

15.2 策略评价报告

在这一节中，作者将针对前文所述的择时策略案例和选股策略案例，给出相应的策略评价报告，以供读者参考。依照上一节所述的策略评价体系，我们主要展示第 14 章 14.2 节中万分之 0.25 手续费时的均线趋势策略和自回归策略，以及第 14 章 14.3 节中的考虑交易成本的股票配置策略。

表 15-1 是仓位优化的均线趋势策略在万分之 0.25 手续费下的模拟交易评价报告，表 15-2 则是仓位优化的自回归策略在万分之 0.25 手续费下的模拟交易评价报告。对比可见，自回归策略的收益率更高，风险水平却相对更低，其风险调整后收益的几项指标也就优于均线趋势策略。

考虑交易成本的股票配置策略则通过表 15-3 ～表 15-5 三个表来展示。其中表 15-3 是单纯的多因子股票配置策略的模拟交易评价报告，表 15-4 是用作对比的同期

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

沪深 300 指数的评价报告。可以看到，相较于沪深 300 指数，多因子股票配置策略的收益更高，同时风险也略大，其风险调整后收益要强于沪深 300 指数。

将考虑交易成本的股票配置策略收益率逐月减去沪深 300 指数收益率，得到新的收益率序列，其评价报告可见表 15-5。可以看到，虽然收益率之差可以提供正收益，但是从风险调整后收益来看，其业绩表现并不十分突出。

需要指明的是，由于择时策略选取的数据是日度数据，决策频率是日度，因此策略净值的更新频率也是日度。而选股策略选取的数据是月度，决策频率是月度，因此策略净值的更新频率就是月度的。

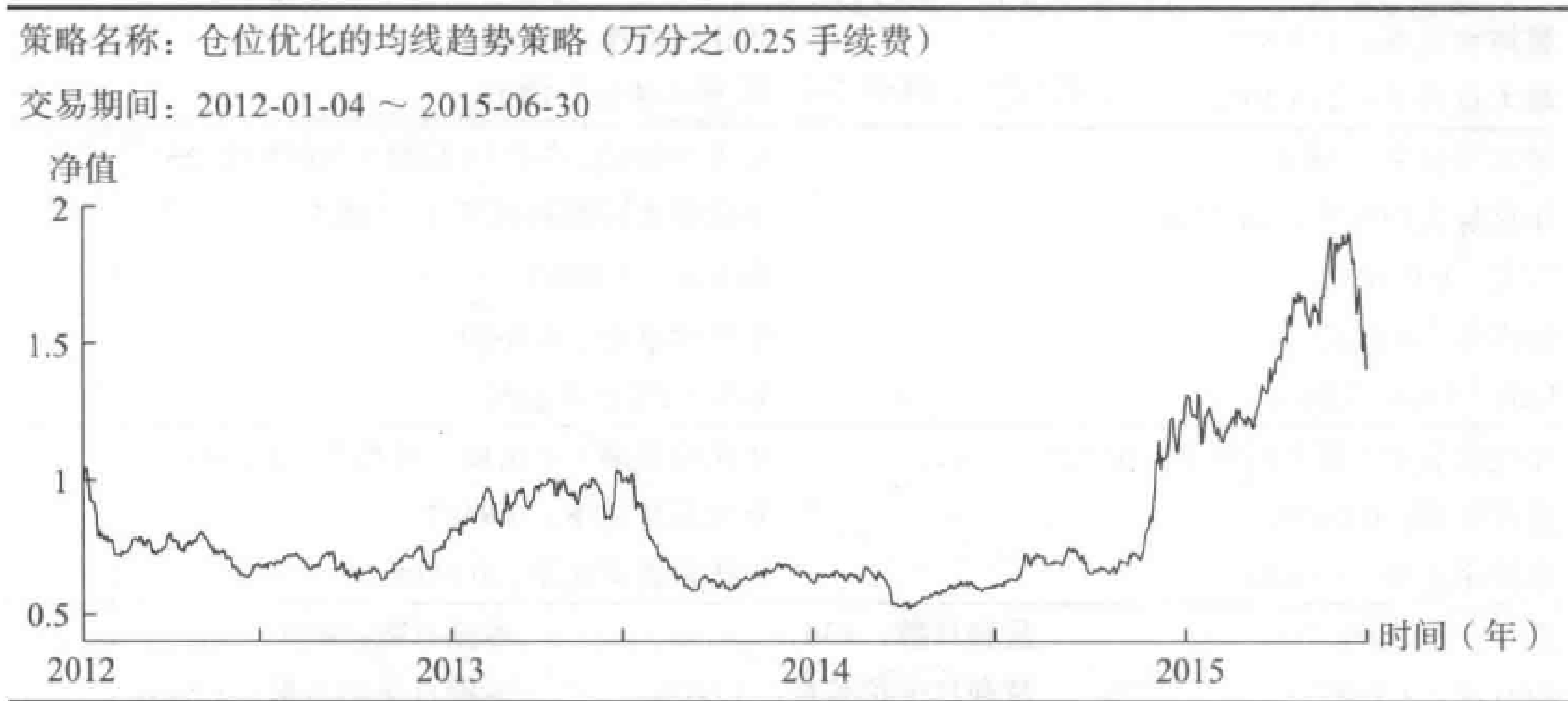
然而我们知道，净值水平实际上是随着交易资产价格的变化而变化的，这种变化随时都在发生。也就是说，不论是基于日度还是基于月度的策略评价，都对真实的净值走势信息进行了压缩。只是很明显，基于月度的策略评价对于信息有更大程度上的压缩，评价的失真程度也就更大。当我们横向对比不同频率下的多个策略时，要特别注意这个问题。

此外，本节中的策略评价报告框架，只是一个我们用来展示的实例。不同的研究人员完全可以根据自己的实际需求、对策略的理解、对评价指标的理解等方面，有针对性地选取对自己更为合适的策略评价内容。例如，可以加入由亏转盈的最长时长、连续亏损的时长和连续盈利的时长，或者在研究择时策略时，基于仓位变化的交易节点给出更多的评价指标，如交易次数，等等。

表 15-1 万分之 0.25 手续费时的均线趋势策略

策略名称：仓位优化的均线趋势策略（万分之 0.25 手续费）

交易期间：2012-01-04 ～ 2015-06-30



全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

（续）

整体收益率：40.36%	年化收益率：10.17%	
最大盈利率：90.58%	最大亏损率：47.82%	
最大回撤率：50.05%	最大回撤期：2012-01-05 ～ 2014-04-09	
年化最大回撤率：30.06%	年化最大回撤时长（日）：74.2	
方差：0.0005	半方差：0.0005	
标准差：0.0230	年化标准差：0.3579	
VaR（5%）：3.72%	VaR（1%）：5.29%	
年化收益率 / 最大回撤率：0.2032	年化收益率 / 年化最大回撤率：0.3384	
夏普比率：0.0238	年化夏普比率：0.2038	
索提诺比率：0.0242	年化索提诺比率：0.2075	
胜率：50.06%	盈利日数：423	亏损日数：422
盈亏比：1.0666	盈利日平均盈利：1.62%	亏损日平均亏损：1.52%

表 15-2 万分之 0.25 手续费时的自回归策略

策略名称：仓位优化的自回归策略（万分之 0.25 手续费）

交易期间：2012-01-04 ～ 2015-06-30

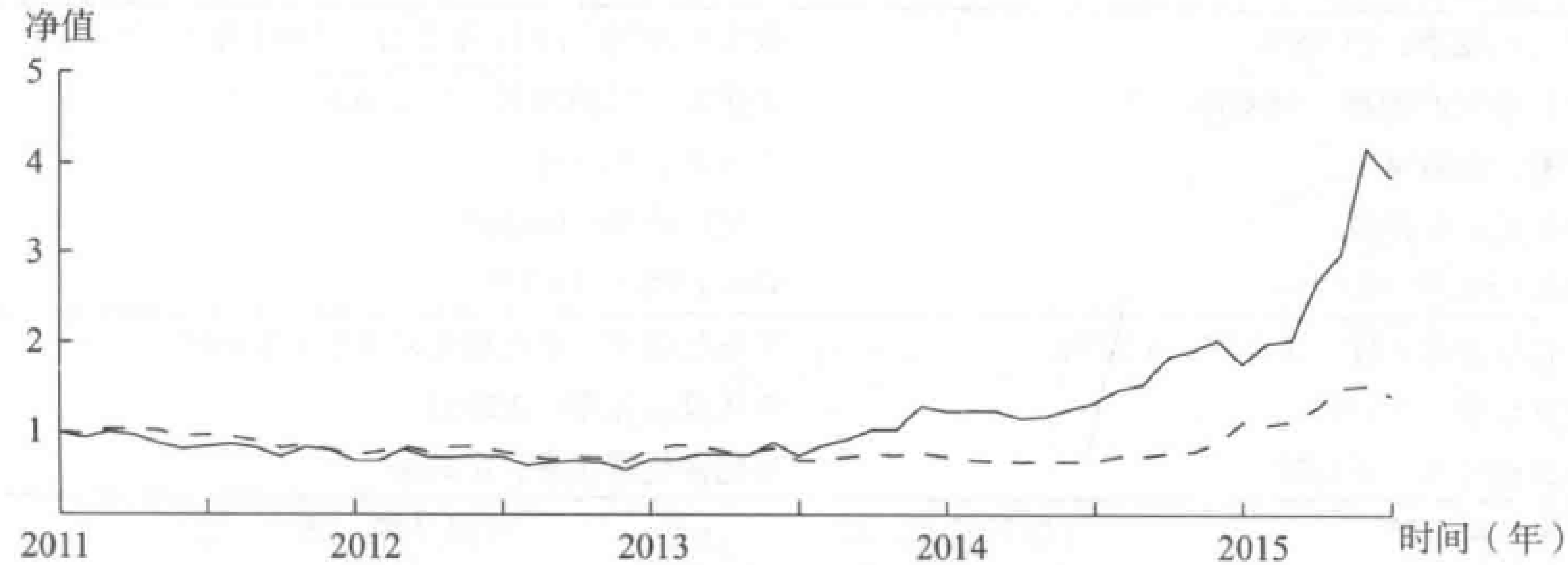
整体收益率：173.87%	年化收益率：33.36%	
最大盈利率：238.19%	最大亏损率：3.62%	
最大回撤率：34.27%	最大回撤期：2014-12-22 ～ 2015-02-25	
年化最大回撤率：21.74%	年化最大回撤时长（日）：43.1	
方差：0.0005	半方差：0.0005	
标准差：0.0232	年化标准差：0.3600	
VaR（5%）：3.66%	VaR（1%）：5.24%	
年化收益率 / 最大回撤率：0.9735	年化收益率 / 年化最大回撤率：1.5343	
夏普比率：0.0579	年化夏普比率：0.8467	
索提诺比率：0.0603	年化索提诺比率：0.8809	
胜率：51.60%	盈利日数：436	亏损日数：409
盈亏比：1.1087	盈利日平均盈利：1.66%	亏损日平均亏损：1.50%

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

表 15-3 考虑交易成本的股票配置策略

策略名称：考虑交易成本的股票配置策略

交易期间：2011 年 1 月～2015 年 6 月



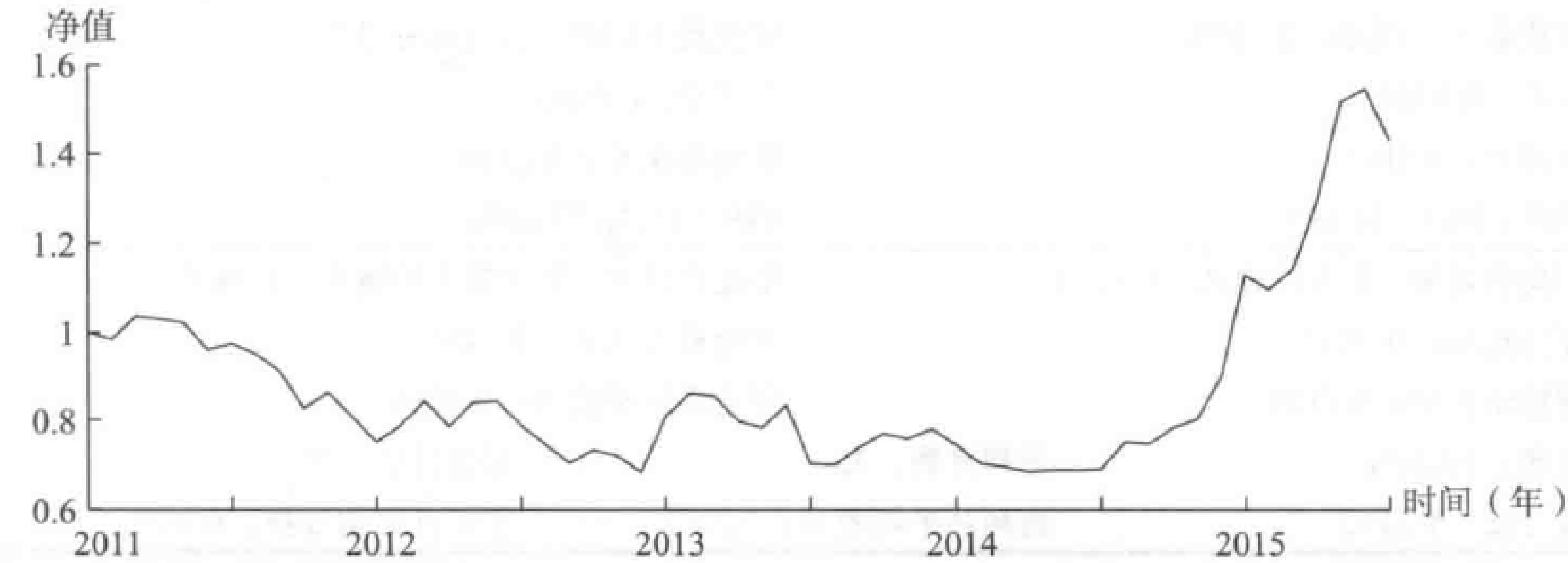
整体收益率：287.18%		年化收益率：35.10%
最大盈利率：319.90%		最大亏损率：40.56%
最大回撤率：40.95%		最大回撤期：2011 年 2 月～2012 年 11 月
年化最大回撤率：18.37%		年化最大回撤时长（月）：4.7
方差：0.0129		半方差：0.0093
标准差：0.1137		年化标准差：0.3940
VaR（5%）：15.58%		VaR（1%）：23.33%
年化收益率 / 最大回撤率：0.8572		年化收益率 / 年化最大回撤率：1.9105
夏普比率：0.2533		年化夏普比率：0.8155
索提诺比率：0.2983		年化索提诺比率：0.9603
信息比率：0.2180		年化信息比率：0.6329
胜率：59.26%	盈利月数：32	亏损月数：22
盈亏比：1.3908	盈利月平均盈利：9.55%	亏损月平均亏损：6.87%

注：净值走势图中实线为策略，虚线为沪深 300 指数；信息比率的基准为沪深 300 指数。

表 15-4 沪深 300 指数（对比用）

策略名称：沪深 300 指数（对比用）

交易期间：2011 年 1 月～2015 年 6 月



全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

（续）

整体收益率：42.99%	年化收益率：8.27%	
最大盈利率：54.75%	最大亏损率：31.60%	
最大回撤率：33.95%	最大回撤期：2011 年 2 月～2012 年 11 月	
年化最大回撤率：16.60%	年化最大回撤时长（月）：6.5	
方差：0.0054	半方差：0.0035	
标准差：0.0732	年化标准差：0.2535	
VaR（5%）：11.12%	VaR（1%）：16.11%	
年化收益率 / 最大回撤率：0.2436	年化收益率 / 年化最大回撤率：0.4982	
夏普比率：0.0913	年化夏普比率：0.2093	
索提诺比率：0.1137	年化索提诺比率：0.2606	
胜率：46.30%	盈利月数：25	亏损月数：29
盈亏比：1.5774	盈利月平均盈利：6.58%	亏损月平均亏损：4.17%

表 15-5 考虑交易成本的股票配置策略收益率减去沪深 300 指数收益率

策略名称：考虑交易成本的股票配置策略收益率逐月减去沪深 300 指数收益率

交易期间：2011 年 1 月～2015 年 6 月

整体收益率：147.00%	年化收益率：22.25%	
最大盈利率：147.48%	最大亏损率：16.51%	
最大回撤率：42.20%	最大回撤期：2014 年 10 月～2014 年 12 月	
年化最大回撤率：15.45%	年化最大回撤时长（月）：3.7	
方差：0.0103	半方差：0.0086	
标准差：0.1015	年化标准差：0.3516	
VaR（5%）：14.48%	VaR（1%）：21.40%	
年化收益率 / 最大回撤率：0.5274	年化收益率 / 年化最大回撤率：1.4407	
夏普比率：0.1937	年化夏普比率：0.5485	
索提诺比率：0.2124	年化索提诺比率：0.6016	
胜率：59.26%	盈利月数：32	亏损月数：22
盈亏比：1.2192	盈利月平均盈利：7.37%	亏损月平均亏损：6.05%

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

第 16 章

结　　语

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

16.1 内容总结

从第 2 章开始，一直到第 15 章结束，就是这本书关于量化交易策略研发的全部内容介绍（见图 16-1）。在这一章中，我们将针对这些内容给出最后的总结，重点在于对这 14 章的内容进行逻辑上的梳理，帮助读者更好地理解作者的思路。随后，我们会讨论这一套研发流程的局限性，以及相应的处理方法。

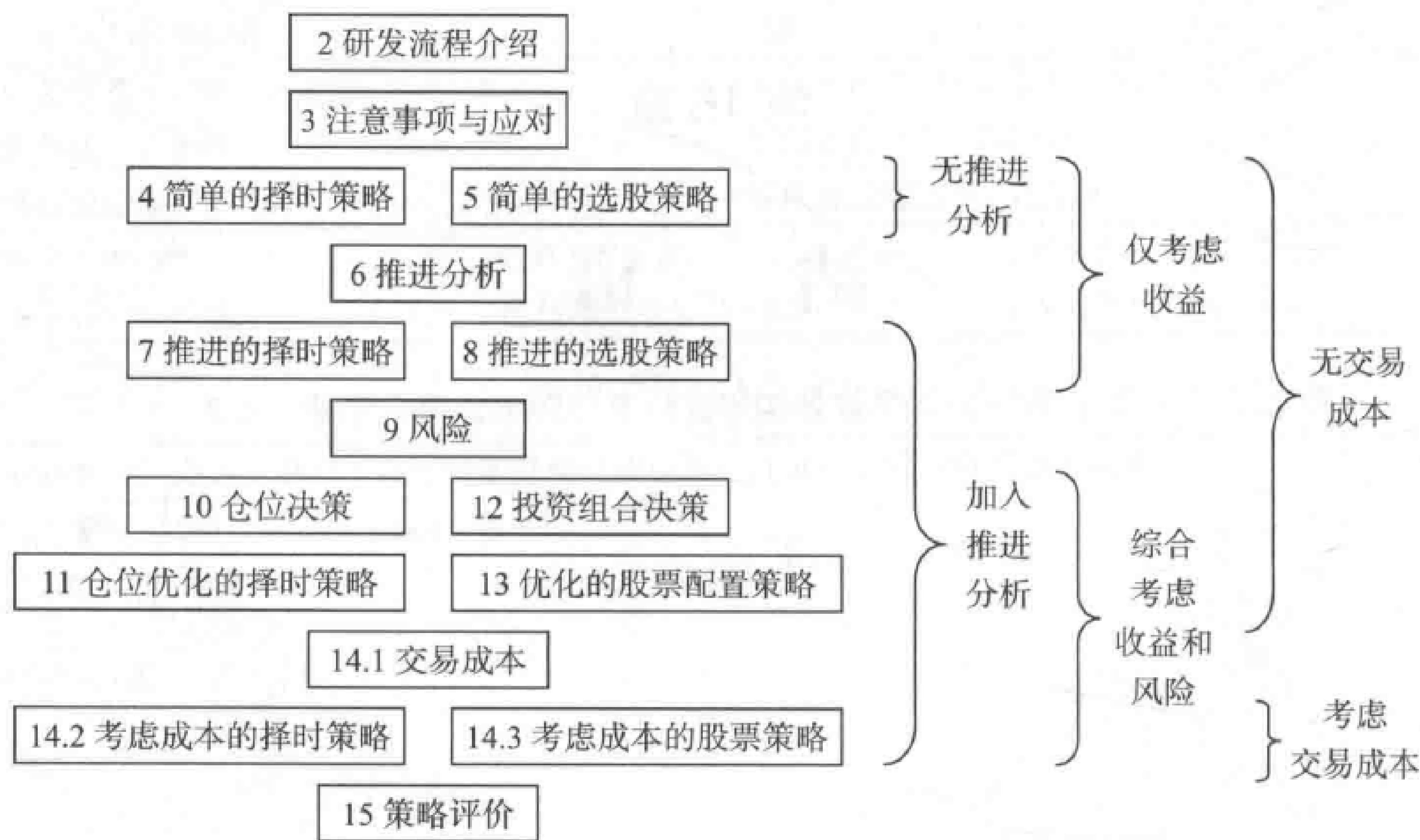


图 16-1 量化交易策略研发的介绍内容

图 16-1 展示了书中关于量化交易策略研发的介绍内容的大致逻辑。可以看到，我们基本上还是遵循着图 2-1 所示的量化交易策略的基本研发流程，来完成整本书的内容介绍。首先是仅考虑收益来形成交易策略，然后同时考虑收益和风险来形成交易策略，最后在所形成策略的基础上考虑交易成本的影响。当然，在这三部分之前，还介绍了不进行推进分析的简单策略，作为起始。

章节的排列顺序，基本上也就是上述思考逻辑的落地结果。最开始的第 2 章，我们对量化交易策略的基本研发流程进行了总览性的介绍。其中 2.1 节是最为基本的策略研发流程，2.2 节则是若干个基本研发流程的变形。后面章节中实际案例所使用的就是 2.1 节的基本流程和 2.2 节中的第一个研发流程。

第 3 章提及了量化交易策略研发中需要注意的若干事项和相应的应对办法。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

大致分为三类：未来信息的规避；过度拟合与欠拟合；回溯测试与真实环境的差异。这是我们在研发量化交易策略时，需要提前做好准备并始终保持注意的几个问题。

在做完第 2 章和第 3 章中的准备工作之后，就开始对基本的交易策略案例进行介绍，具体包括第 4 章中简单的择时策略和第 5 章中简单的选股策略。称其为“简单”，一是因为策略只根据收益判断买卖、不做仓位决策；二是因为策略使用全样本数据进行简单优化，这样会导致未来信息的问题。虽然简单且存有隐患，但是却是适合读者进行初步认知的案例。

其中第 4 章中简单的择时策略，分为均线趋势策略、均线反转策略和自回归策略。均线策略是分类型择时策略的典型案例，而自回归策略则是预测型择时策略的典型案例。这一章中的均线反转策略，虽然存在正收益，但是在网格搜索的过程中已经表现出一定的问题。

第 5 章中简单的选股策略，也被称为因子选股策略。这一章介绍了两个因子，即市值因子和反转因子，以及相应的单因子选股策略。将两个因子通过模型组合起来并用于选股，就形成了 5.4 节中的多因子选股策略。

可以看到，第 4 章和第 5 章在基本一致的前提条件下，分别介绍了量化择时策略和量化选股策略，这也是后面的内容将一直保持的撰写风格。实际上，作者也希望通过这样的处理来表达自己的一个想法，即择时策略和选股策略虽然一个偏向于时间序列、一个偏向于横截面，但是研究过程都是由收益、风险所驱动的优化行为，在交易策略的本质上两者是相近的。

第 6 章介绍了推进分析，这是一种可以使回溯测试免受未来信息影响的样本外检验方法。其操作流程非常贴合实际交易的情况，因此适合在研究量化交易策略时使用，方便得到更值得信任的策略回溯测试结果。除此之外，我们还介绍了多层推进分析和推进分析下的验证，用以拓展读者的思路。

第 7 章和第 8 章是在回溯测试过程中使用了推进分析的择时策略案例和选股策略案例。由于推进分析的引入，均线策略还能够在趋势和反转之外增加所谓的混合策略。实际上，通过推进分析，我们也确认了均线反转策略在该样本下的亏损实质。此外，我们还展示了自回归策略和多因子选股策略的多层推进分析结

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

果，而多层推进分析的结果则进一步支持了这两个策略在所选样本上的有效性。

第 9 章论述了在量化交易策略的研究与评判中所使用到的风险概念，重点在于对一些风险度量指标的介绍，具体包括常用的风险度量、其他风险度量以及风险和收益的结合等。偏向于交易层面的止损也在这一章中有所涉及。

介绍了风险的概念之后，就可以给出在收益和风险共同作用下，优化了买卖点和仓位的择时策略，以及优化了股票选择和仓位配比的股票配置策略。两者又各分两章进行介绍，前一章重点在于对决策模型和风险理念的阐述，后一章重点在于案例分析。当然，这部分的内容也是在推进分析的框架下完成的。

择时策略的内容在第 10 章和第 11 章。第 10 章介绍了做仓位决策时会使用到的模型，重点在于 10.1 节中的凯利公式。同时，针对凯利公式得出的仓位结果往往会造成风险过大的问题，给出了作者认知下的实用的仓位决策方法。第 11 章就是进行了仓位决策之后的择时策略案例介绍，包括仓位优化的均线趋势策略和仓位优化的自回归策略。

股票配置策略的内容在第 12 章和第 13 章。第 12 章介绍了确定个股选择后具体的仓位配置模型，也就是我们常说的投资组合决策模型。一个典型且基础的工作是最优投资组合理论，在其基础上，我们还介绍了若干个其他形式的投资组合优化方法，来克服基本的最优投资组合理论所具有的缺陷，特别是实际使用时的缺陷。

第 13 章的 13.1 节其实仍然是一个改进的投资组合决策模型，该模型基于多因子的思路重构了协方差矩阵的估计过程，较为适用于本书所举的选股策略的案例。13.2 节则介绍了使用该模型配置股票仓位之后，交易策略具体的回溯测试结果，其业绩表现与普通的等权重组合非常相近。

第 14 章考察了交易成本的构成，以及交易成本对前面所举案例的业绩影响。14.1 节简单介绍了两大类交易成本，即显性交易成本和隐性交易成本。其中显性交易成本包括交易税费和买卖价差，隐性交易成本包括滑点和冲击成本。在考虑最为简单的交易税费的基础上，14.2 节和 14.3 节分别给出了择时策略和股票配置策略的业绩表现情况，这是我们最后可以用于判断的策略表现结果。

第 15 章是独立于第 2 章给出的完整研究流程之外的一部分内容，我们总结

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

了对量化交易策略进行评价时需要包含的内容，并给出了第 14 章中若干个策略的模拟交易评价报告，供读者参考。

16.2 研发流程的局限与应对

16.2.1 研发流程的局限

第 1 章 1.2 节论述了量化交易策略的一些缺陷，包括易复制、易暴露、无法兼容非量化信息、转向缓慢等。可以看到，这些缺点都是量化交易策略所固有的，当我们选择使用量化交易这一策略门类时，实际上已经默认地接受了这些缺陷及其所带来的危害。

与此同时，量化交易策略还存在着一个本质上的缺陷，那就是过度拟合所造成的策略回溯测试结果偏离实际的问题。不同的是，其与量化交易策略的整个研发流程息息相关，因此作者选择整体介绍完量化交易策略研发之后再对此进行阐述。当然，将这部分内容放置在最后一章，也是出于着重强调该问题的目的。

在第 3 章 3.2 节中，作者曾介绍过量化交易策略研发中过度拟合和欠拟合的来源情况。如图 16-2 所示，两个圆分别代表数据内部特征和交易策略，两者重合的浅灰色区域是策略挖掘出数据特征进而获利的区域，深灰色区域代表欠拟合的部分，白色区域则是策略无效，即过度拟合的部分。

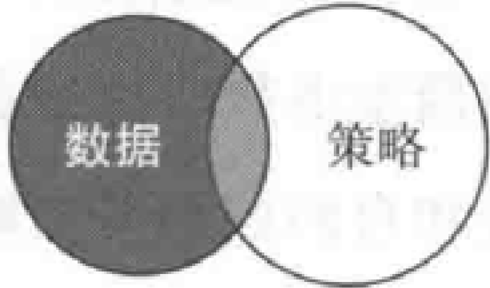


图 16-2 过度拟合与欠拟合

实际上，欠拟合的深灰色部分并不会对策略的可行性造成实质影响，因为我们研发量化交易策略的目的并不是全知全能，而只是依靠浅灰色的生效部分获得盈利。但策略部分超出数据特征的白色区域则不同，由于研发过程中优化方法的引入，过度拟合往往会不切实际地提升回溯测试的盈利结果，从而造成研究人员的误判。我们在案例中使用到的参数的网格搜索、回归式的拟合等，都是最优化的手段。

具体来说，在研发量化交易策略时，我们可能会优化收益、优化风险，也可能会针对风险调整后收益进行优化。当存在过度拟合的问题时，我们就会相应地

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

得到过高的收益和过低的风险。不管是哪种情况，最后供我们参考的结果都会比实际应用策略时要好，而这种“好”是虚假的。

我们在第 3 章中曾经提到，推进分析是一种处理过度拟合问题时行之有效的分析技巧，通过推进分析等处理，似乎是将这些优化手段所造成的过度拟合问题规避了。但是遗憾的是，不论是推进分析还是多层推进分析，都只能模拟实际交易过程，并提供模拟环境下的样本外检验，而无法真正地解决过度拟合的问题。根源在于，无论我们如何划分样本数据，并做出各种各样的检验，在研究过程中所有的样本数据对我们来说都是可见的。真正不可见的数据只可能是未来的数据，如图 16-3 所示。因此，真正意义上的样本外检验，也只可能是研发确定交易策略之后的实际操作或实际检验。

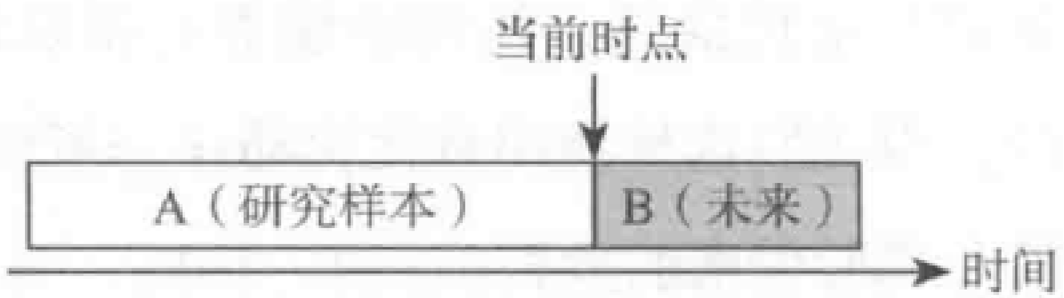


图 16-3 未来信息示意图

除此之外，推进分析只能针对形式明确的最优化方法来规避过度拟合问题。对于那些不能使用数量化手段进行表述的优化方法，推进分析之类的处理手段非常难以实施，仅存在理论上的可行性。

例如，在给出市值因子和反转因子时，我们已经试验过了很多没有显著作用的因子。择时方面，我们也尝试了很多种不同的时间序列模型，才最终选定了均线和自回归两种策略进行介绍。很明显，这些没有展示出来但是确实存在的研究行为，实际上可以归为隐性的最优化处理，也是完全有可能造成过度拟合问题的。

甚至可以说，不论我们使用多么精妙的技术手段来进行处理，只要有“研究”本身的存在，我们就不可能在量化交易策略研发的过程中完全规避优化行为，也就会因此受到过度拟合的困扰。

此外，国内的资本市场运行时间尚短，存在着数据量过少、市场结构变化过快等问题。当我们使用这些数据作为整体样本进行研究时，就会进一步地加深最优化所带来的过度拟合问题。

这样看来，通过一般研发流程得到的量化交易策略，其回溯测试结果都可能会被或多或少的高估。如果一个实际上并不具备显著盈利能力的策略，通过“研

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

发”流程而被选中并进入实际运行，其结果就是资金的浪费，同时承担了无谓的风险。鉴于我们可以将这部分资金用于其他可盈利的策略或项目上，可以认为我们损失了所谓的机会成本。

而一个策略如果被低估了风险，按照本书所介绍的基本研发流程来看，风险直接驱动仓位决策，结果往往就是在实际运行中使用的仓位过重。在第 10 章 10.1 节论述凯利公式时曾有涉及，如果仓位较之真实的最优仓位更重，会导致收益下降甚至产生损失。更重要的是，对风险的低估会导致策略使用者在策略运行时无法准确把握净值的实际波动，甚至在不恰当的时机提前终止策略。

16.2.2 主观逻辑与经济直觉

要在技术上应对这一问题，其实并没有非常有效的方法。研究人员在熟悉量化交易策略的研发流程之后，对过度拟合的程度进行人为判断，自然是可行的办法之一。但是这样的判断缺少系统性和准确性，有效程度值得怀疑。

作者较为推荐的做法是，使用主观逻辑的支持来弥补对于回溯测试结果的信心缺失。更为学术一点的说法是，让量化交易策略符合“经济直觉”。当策略本身可以体现出经济上的逻辑或者事物内在的逻辑时，使用者就更有理由相信，策略在回溯测试下之所以盈利是因为挖掘出了历史数据中的本质特征，而非纯粹的过度拟合。

例如人们会根据自己的理解，在逻辑上对金融资产的趋势特性进行解释。大部分的解释归因于市场参与者方面，认为是人性的贪婪和恐惧造成的资产价格上的趋势，或者趋势来自市场参与者对于信息的反馈速度不一致，等等。

市值因子的经济逻辑相对而言更为清晰，在经济水平上升的现实情况下，市值较小的公司往往更有发展潜力，从整体上也就体现出更强的盈利能力。当然，也有一些人从市值小的股票更容易操控的角度来理解这个问题。

反转因子的逻辑在解释上与趋势特征是同源的，市场参与者的个体特征经过市场的融合，最终形成了资产价格走势的反转特性。只不过对比趋势而言，反转的作用时间往往较短，并造成了相反的结果。

当然，上述经济逻辑，有可能只是研发人员通过纯粹数量化的研究得到了趋

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

势、市值、反转这些有效策略之后，再根据策略形式构造出的内部逻辑，在顺序上逻辑是处于量化研究之后的。更令人信服的，应该是先基于经济逻辑构造出相应的量化交易策略，然后回溯测试发现其确实有效，这或许是对研究人员来说最为理想的一种情况。

另一方面，这里所说的主观逻辑或者经济直觉，与完全人为判断的主观交易又有不同。人的逻辑思维有些是可以通过数量模型来描述的，我们要利用的也正是这一部分的逻辑思维。此时的逻辑只是用来辅助和支持数量模型的，量化模型和量化决策过程仍然是我们处理的主体，这与本书最开始给出的量化交易策略的定义没有冲突。

而随着技术的不断发展，也许在未来的某一个时刻，量化交易策略将可以在数据挖掘的基础上，自主地形成另一种形式上的“意义”，进而将逻辑思维和量化分析合二为一。我们已经看到了数据科学、人工智能等领域中让人赞叹的成果，也理应在技术进步的面前保持谦卑，持续学习。

最后需要说明的是，本书的重点在于介绍一套用于研发量化交易策略的流程和技术，案例中的若干策略其实都非常简单。而在实际的策略研发中，往往会运用到一些更为复杂的模型和策略。研究人员有必要理解这些模型，知道其意义何在，才有可能正确地使用它们，乃至跳出基本模型的限制，尝试构造更为适合自己的策略。量化交易策略本身的数理逻辑是应用这些策略的前提和基础，这需要我们在实际的工作中进行掌握。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

参考文献

[1] 爱德温·李费佛.股票作手回忆录[M].真如，译.海口：海南出版社，2008.

[2] 安德鲁·波尔.统计套利[M].陈雄兵，张海珊，译.北京：机械工业出版社，2011.

[3] 柯蒂斯·费思.海龟交易法则[M].乔江涛，译.北京：中信出版社，2013.

[4] 里什·纳兰.打开量化投资的黑箱[M].郭剑光，译.北京：机械工业出版社，2012.

[5] 理查德·格林诺德，雷诺德·卡恩.主动投资组合管理[M].李腾，杨柯敏，刘震，译.北京：机械工业出版社，2014.

[6] 罗伯特·帕多.交易策略评估与最佳化[M].李佳儒，译.台北：寰宇出版社，2010.

[7] 罗杰·洛温斯坦.赌金者：长期资本管理公司的升腾与陨落[M].毕崇毅，译.北京：机械工业出版社，2017.

[8] 欧内斯特·陈.量化交易：如何建立自己的算法交易事业[M].黄嵩，译.大连：东北财经大学出版社，2014.

[9] 欧内斯特·陈.算法交易：制胜策略与原理[M].高闻酉，黄蕊，译.北京：机械工业出版社，2017.

[10] ASNESS C, MOSKOWITZ T, PEDERSEN L. Value and momentum everywhere[J]. The Journal of Finance, 2013, 68(3): 929-985.

[11] ASNESS C. Variables that explain stock returns: simulated and empirical evidence[D]. University of Chicago, 1994.

[12] BARRA. United States equity risk model handbook version 3 (USE3)[R]. BARRA Inc., 1998.

[13] BLACK F, LITTERMAN R. Global portfolio optimization[J]. Financial Analysts Journal, 1992, 48(5): 28-43.

[14] FABOZZI F, FOCARDI S, KOLM P. Trends in quantitative finance[M]. Charlottesville: CFA Institute Research Foundation, 2006.

[15] FAMA E, MACBETH J. Risk, return and equilibrium: empirical tests[J]. Journal of Political Economy, 1973, 81(3): 607-636.

[16] FAMA E, FRENCH K. Common risk factors in the returns on stocks and bonds[J]. Journal of Financial Economics, 1993, 33(1): 3-56.

[17] FAMA E, FRENCH K. The cross-section of expected stock returns[J]. The Journal of Finance, 1992, 47(2): 427-465.

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

[18] FAMA E. Efficient capital markets: a review of theory and empirical work[J]. The Journal of Finance, 1970, 25(2): 383-417.

[19] KELLY J. A new interpretation of information rate[J]. Bell System Technical Journal, 1956, 35(4): 917–926.

[20] LO A, MAMAYSKY H, WANG J. Foundations of technical analysis: computational algorithms, statistical inference, and empirical implementation[J]. The Journal of Finance, 2000, 55(4): 1705-1765.

[21] MARKOWITZ H. Portfolio selection[J]. The Journal of Finance, 1952, 7(1): 77-91.

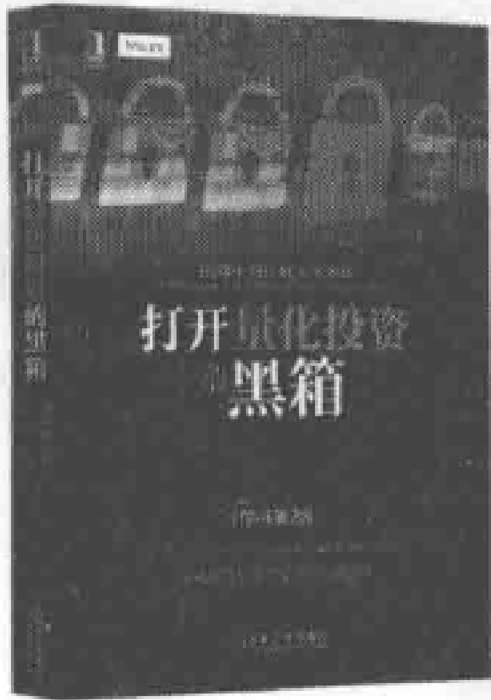
[22] ROSS S. The arbitrage theory of capital asset pricing[J]. Journal of Economic Theory, 1976, 13(3): 341-360.

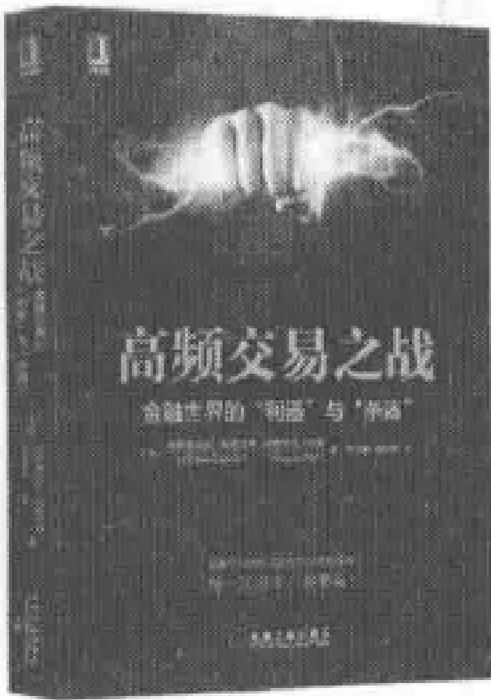
[23] SHARPE W. Capital asset prices: a theory of market equilibrium under conditions of risk[J]. The Journal of Finance, 1964, 19(3): 425-442.

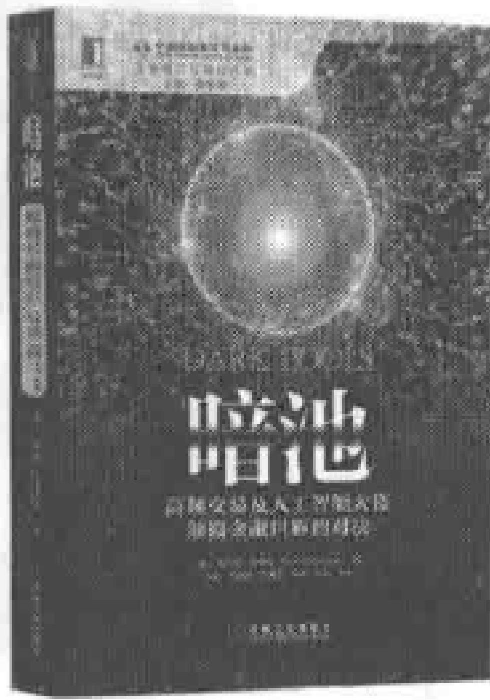
[24] VINCE R. The mathematics of money management: risk analysis techniques for traders[M]. Hoboken: Wiley, 1992.

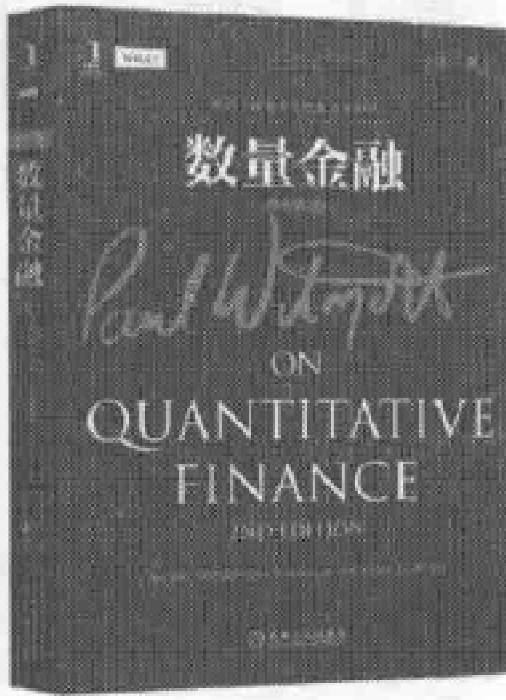
全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

推荐阅读









书名	作者	ISBN	价格
透视高频交易	林建	978-7-111-50372-9	49.00
高频交易之战：金融世界的“利器”与“杀器”	（法）弗雷德里克·勒雷艾弗、弗朗索瓦·比雷	978-7-111-50565-5	45.00
暗池：高频交易及人工智能大盗颠覆金融世界的对决	（美）斯科特·帕特森	978-7-111-51299-8	59.00
数量金融（原书第2版）（第1卷）	（美）保罗·威尔莫特	978-7-111-49501-7	119.00
数量金融（原书第2版）（第2卷）	（美）保罗·威尔莫特	978-7-111-49487-4	119.00
数量金融（原书第2版）（第3卷）	（美）保罗·威尔莫特	978-7-111-49500-0	149.00
打开高频交易的黑箱	（美）迈克尔·德宾	978-7-111-44658-3	40.00
打开量化投资的黑箱（原书第2版）	（美）里什·纳兰	978-7-111-53729-8	69.00
高频交易	（美）里什·纳兰	978-7-111-37321-6	45.00
主动投资组合管理：创造高收益并控制风险的量化投资方法（原书第2版）	（美）理查德 C. 格林德、雷诺德 N.卡恩	978-7-111-47472-2	100.00
可视化量化金融	（美）迈克尔·洛夫雷迪	978-7-111-48758-6	59.00

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

蒋炜航
宜信大数据创新中心技术副总经理 宜信指旺财富总经理

《量化炼金术》是“量化交易”这一引人遐想的主题下不可多得的读物。基于以下几点，我向身边的多位朋友推荐了本书：首先，本书的形式是体系化的，这对帮助大家理解“量化交易”这样一个有一定门槛的主题非常有效；其次，本书针对各个知识点都附带了描述清晰的实例，对于想动手实践文中内容的读者很有帮助；最后，本书介绍相关内容的出发点是让读者“知其然也知其所以然”，我相信这样才是从事量化研究以及其他领域研究的正确方式。因此，在这里我也向大家郑重地推荐本书，希望大家和我一样享受阅读。

梁新刚
爱智慧科技创始人&CEO 百米生活股份有限公司前CEO

本书循序渐进地构建了一套极具操作性的量化交易策略的开发流程和评估优化体系。

中国股市正处于结构性大变革，量化交易比重势必提升。本书对于好奇的散户、乐在其中的宽客和专业投资机构都很有参考价值！

龚朴
华中科技大学管理学院财务金融系教授、博士生导师

这是一本思路清晰、案例丰富、可读性强的著作，作者通过介绍量化交易策略的研发流程，将量化投资领域里的多个经典模型与技术串联了起来，其间也不乏作者自己的创新之处和实践经验，更难能可贵的是提炼出了一套自己的量化思维模式，值得反复阅读。



上架指导：金融投资

ISBN 978-7-111-57532-0

9 787111 575320

定价：59.00元

投稿热线：(010) 88379007
客服热线：(010) 68995261 88361066
购书热线：(010) 68326294 88379649 68995259

华章网站：www.hzbook.com
网上购书：www.china-pub.com
数字阅读：www.hzmedia.com.cn